REPORT RESUMES

ED 010 649 24

AN INVESTIGATION OF MARKING PRACTICES OF SECONDARY SCHOOL TEACHERS.

BY- TERWILLIGER, JAMES S.
GEORGE PEABODY COLL. FOR TEACHERS, NASHVILLE,TENN.
REPORT NUMBER CRF-S-495 PUB DATE AUG 66
REPORT NUMBER BR-5-8342
CONTRACT OEC-S-619-65
EDRS PRICE MF-$0.18 HC-$4.32 108P.

DESCRIPTORS- *GRADING, *SCORING, *RATING SCALES, SECONDARY SCHOOLS, STATISTICAL DATA, *INSERVICE TEACHER EDUCATION, NASHVILLE, TENNESSEE

VARIOUS ASPECTS OF THE MARKING PRACTICES OF 39 SECONDARY SCHOOL TEACHERS FROM TWO SCHOOLS IN METROPOLITAN NASHVILLE-DAVIDSON COUNTY, TENNESSEE, WERE STUDIED. SPECIAL MARKING EXERCISES CONTAINING STANDARD DATA ON HYPOTHETICAL STUDENTS WERE USED TO STUDY THE MARKING PRACTICES OF TEACHERS UNDER MORE UNIFORM CONDITIONS THAN EXIST IN THE CLASSROOM. A COMPARISON OF THE MARKING PRACTICES DURING ONE SEMESTER WITH THE PRACTICES DURING THE FOLLOWING SEMESTER, AFTER AN INSERVICE TRAINING PROGRAM, SHOWED THE INSERVICE TRAINING HAD LITTLE EFFECT ON THE TEACHERS' GRADING PRACTICES AND NO IMPACT ON THE MARKING EXERCISES. THE STUDY SHOWED THERE IS GREAT VARIABILITY IN THE AVERAGE MARKS OF INDIVIDUAL TEACHERS ON ALL TYPES OF CLASSROOM MEASURES. THE MARKING EXERCISES ALSO REVEALED THAT MOST TEACHERS INTERPRET NUMERIC DATA IN AN ABSOLUTE RATHER THAN A NORMATIVE FASHION. (AL)

ED010649

(5-8342)
(S-495)

RECEIVED
RESEARCH BRANCH
SEP 12 1966
DIVISION OF
Elementary-Secondary Research

NAA

AN INVESTIGATION OF MARKING PRACTICES OF SECONDARY SCHOOL TEACHERS

James S. Terwilliger

GEORGE PEABODY COLLEGE
FOR TEACHERS

Final Report of a Project Supported
by the U. S. Office of Education
Under Contract S-619-65.

August, 1966

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

AN INVESTIGATION OF MARKING PRACTICES OF SECONDARY SCHOOL TEACHERS

JAMES S.TERWILLIGER

GEORGE PEABODY COLLEGE
FOR TEACHERS

AUGUST, 1966

Final report of a project supported by the U.S. Office of Education under contract S-619-65. Reproduction, in part or in whole, for the use of the U.S. Government is permitted.

Acknowledgements

The author wishes to express his appreciation to the principals and teachers of the two schools used in this study. Their cooperation and interest in the project greatly facilitated all phases of the research.

A special word of gratitude is also due to Joyce Lytle, Jane Weyer, Richard Hooper, and James Tillman, all of whom served as research assistants on the project. Their dedication and efforts in the collection and analysis of the data contributed greatly to the study.

Abstract

Various aspects of the marking practices of 39 secondary school teachers from two schools in Metropolitan Nashville-Davidson County, Tennessee were studied intensively. The impact of a special in-service program on the assignment of marks was also investigated.

Analysis of marks assigned by the teachers during the first semester of 1965-66 showed that average marks on quizzes and daily work are higher than those assigned for six-week tests and that marks assigned for six-week periods are higher than those on semester exams. Despite these mean differences, there is great variability in the average marks of individual teachers on all types of classroom measures. The same general results were obtained for second semester data (following the in-service program) with the only difference being that average marks for _all_ types of classroom measures are lower for the second semester.

A positive correlation was found between average class marks and average class IQ for both semesters. Correlations between IQ and marks assigned by individual teachers vary from slightly negative to above .80. Individual differences in the magnitude of these correlations can be only partially accounted for by differences in range of ability from one class to the next. The relationships among marks for different six-week periods and the final semester exam tend to be uniformly high irrespective of the magnitude of the relationship between IQ and semester marks. This is true of the data for both semesters.

Variation in final semester marks is highly predictable from marks assigned for six-week periods and the semester exam. The semester exam and third six-week marks generally account for more variance than do marks for the first two six-week periods. However, there is great variation from teacher to teacher in the relative importance of six week and semester exam marks.

Special Marking Exercises containing standard data on hypothetical students were used to study the marking practices of teachers under more uniform conditions than exist in the classroom. Inter-teacher agreement is generally high on these exercises although there is great variation in the marks assigned to specific students. The variance in marks assigned on these exercises is almost entirely accounted for by data in the exercises but the importance of specific types of data tend to be determined largely by the statistical characteristics of the scores built into the exercises, e.g. a quiz with large variability is more important than a major exam. This strongly suggests that teachers are largely unaware of the effects of differential variability in data with which they work. Also, the marking exercises reveal that most teachers interpret numeric data in an absolute rather than normative fashion. The in-service training had no apparent impact upon the marking practices on the exercises.

Finally, the self-reported marking practices of the teachers reflect great individual differences with respect to the basis for marking and the importance attached to a multitude of factors considered in the assignment of marks. No substantial changes in self-reported practices resulted from the in-service program.

INTRODUCTION

The assignment of marks to students is one of the most important responsibilities of the classroom teacher. The marks which a student receives influence his life in several ways. First, marks are typically the only data which a student and his parents regularly receive regarding his success in the school setting. Marks are therefore the primary channal through which the student and his parents have an opportunity to assess the correspondence between actual performance and self- or parental-imposed aspirations. Second, marks are frequently employed for administrative action such as sectioning students according to ability, selecting students for special recognition, etc. The marks a student receives can thereby be decisive in determining the nature of the future training and opportunities which he receives within the school program. Finally, the marks a student receives have a tremendous bearing upon his future opportunities beyond the school setting. School marks have traditionally been one of the most important factors in college admissions and in the selection of personnel for various types of specialized training or employment. The use of marks for such purposes is becoming increasingly important as more and more students seek admissions to college and more and more employers require a secondary education of job applicants. There can be little doubt that marks are of considerable practical significance to a student when one recognizes the variety of ways in which the student's life is influenced by them.

In spite of the almost universal use of marks in our schools and the crucial role which they play in shaping a student's future, relatively little research has been conducted recently on the actual process of assigning and reporting marks. The evidence which has been collected consistently indicates that practices related to the assignment and reporting of marks are better characterized by diversity from one teacher to the next and from one school to the next than by any standardization of practices. See Roelfs (1948; 1955), Jansen (1960) and Vroman, et. al. (1962). This general result holds for individual teachers within a specific subject matter and school as well as for teachers working in different subject matters and different schools. Policies designed to decrease the diversity of marking practices within a school have also been found to vary greatly from school to school. See Jansen (1960) and Terwilliger (1966). The upshot is that marks assigned by one teacher carry quite different meaning than do those assigned by another teacher, even if both teachers teach the same subject at the same school. This is clearly illustrated in the results of a recent survey of the marking practices of nearly 4000 secondary school teachers conducted by the writer. (Terwilliger, 1966)

These data, along with studies of the reliability of marks Starch and Elliot (1912; 1913); Johnson (1925); Morlan (1931); Diederich, et. al. (1961) present a rather dismal picture of the process of marking as it is typically done in the schools. In spite of this, little research has been conducted recently on the marking process. Just as significant as the lack of research is the manner in which the topic of marking is treated in textbooks on measurement and evaluation for teachers. An examination of several of the current texts reveals that only a few pages are typically devoted to marking and the topic is generally discussed from a very broad philosophical point of view. Little is offered in the way of concrete suggestions or specific criticisms of methods currently used by many classroom teachers.

This study attempts to answer several questions concerning the marking practices of secondary school teachers. These questions may be summarized as follows:

(a) With respect to data from quizzes, tests, homework, special projects, etc.

    (1) what is the difficulty level (mean score) and to what extent does this difficulty vary from teacher to teacher?

    (2) what is the discriminability (standard deviation) and to what extent does this vary from teacher to teacher?

    (3) what is the relative weight of each type of data (determined by multiple regression) in the assignment of marks and to what extent do these weights vary from teacher to teacher?

    (4) what is the relationship of standardized test scores to these data and how does this vary from teacher to teacher?

(b) With respect to final semester marks

    (1) what is the average mark assigned and how does this vary from teacher to teacher?

    (2) what is the variability in assigned marks and how does this vary from teacher to teacher?

    (3) what is the relative weight given to marks in previous marking periods and the final semester exam?

    (4) what is the relationship of standardized test scores to semester marks and how does it vary from teacher to teacher?

(c) With respect to the marks assigned to a class of 20 hypothetical students for which certain standard data are provided

(1) what is the reliability of the marks (inter-teacher agreement) and to what extent does this vary from one subject area to the next?

(2) what is the relative weight of various types of data (quizzes, tests, homework, etc.) and how do these weights vary from teacher to teacher?

(d) An analysis of self-reported marking practices will also be made to check for systematic differences between subject matter areas and to compare data collected in this fashion with actual marking data based upon school records and the class of hypothetical students.

(e) A final major question of the study deals with the degree to which the marking practices of secondary teachers can be influenced by a concentrated in-service training program. This question will be answered by comparing actual school marks, marks assigned to hypothetical students, and self-reported marking practices after training with those obtained prior to the training, i.e. a repetition of the analyses described in (a), (b), (c), and (d) after special training.

GENERAL DESIGN

Participating Schools

The project described in this report was undertaken to study intensively the process of student marking in two secondary schools in Metropolitan Nashville--Davidson County, Tennessee. The schools participating in this research were selected so as to represent different geographic and socio-economic cross-sections of the entire Metropolitan Nashville School System. Selected data describing the two schools are presented in Table 1.

Table 1

Selected Descriptive Data on the Two Participating Schools

	School A	School B
Grade Levels	7-12	7-12
Approximate Total Enrollment	1500	1600
Total Faculty	61	67
Age of Building in Years	6	36
Percent of 1966 Graduates College-Bound	85	55
Socio-Economic Characteristics of District Served	Middle to upper class, predominantly upper-middle; father occupation primarily managerial or professional; newer suburban area	Middle to upper-middle, predominantly middle; father occupation primarily managerial or small business; older established residential area
Number of Participating Teachers	21	18

The principals of the two schools were contacted during the latter half of the first semester of the 1965-66 school year. At that time, the general objectives of the research were discussed and the proposed study was outlined in detail. Each principal, in turn, discussed the project with members of his faculty and obtained their cooperation as participants in the study.

Participating Teachers

The teachers participating in the project represent all faculty members in the two selected schools who were teaching "academic" subjects in grade levels 9-12 during the 1965-66 school year. ("Academic" subjects are defined as classes in foreign languages, language arts, mathematics, science, or social studies for purposes of this investigation.) A total of 39 (21 at School A and 18 at School B) participated in the study. Selected biographical data for the participants is presented in Table 2.

The reason for restricting participation to teachers in only the academic areas is simply that the process of marking itself is undoubtedly quite different in academic, vocational, and service courses. Since the objectives of instruction and the methods of evaluation tend to differ for academic as opposed to non-academic subjects, the present investigation was limited to a consideration of only the former. The in-service program described below could therefore be developed with primary focus upon those problems and issues in the assignment of marks which are of most concern in academic subject areas.

Table 2

Biographical Data for the Participating Teachers

	School A	School B	Total
Sex			
Female	16	12	28
Male	5	6	11
Total Years Experience			
Less than 5	5	1	6
6-10	8	5	13
11-20	5	3	8
More than 20	3	9	12
Highest Degree			
B.A. or B.S.	8	6	14
M.A.	12	12	24
Ed.S.	1	0	1
Subject Area			
Foreign Lang.	5	2	7
Lang. Arts	5	4	9
Mathematics	4	3	7
Science	3	3	6
Social Studies	4	6	10

In-Service Program

A special in-service training program on marking practices was developed as part of the over-all study. The primary objective of this program was to discuss with the participating teachers certain philosophical and practical issues involved in the assignment of marks to students.

Each teacher was assigned an identification number at the first training session and was assured that all data collected during the course of the project would be identified by number only. Teachers at School A were assigned the numbers 01-21 and teachers at School B were assigned the numbers 30-43 and 45-48 (Teacher 44 chose not to participate after the first session). These numbers will be used in the presentation and discussion of all data for individual teachers.

The early part of the in-service program was devoted to basic notions in classroom testing and the statistical treatment of test results. The latter portion of the program was concerned with ways in which test and other types of classroom performance data can most appropriately be translated into school marks. An outline of the topics covered in the program and a fuller discussion of the materials used can be found in Appendix A.

The in-service program began during the third week of the second semester of the 1965-66 school year. The program consisted of weekly one-hour sessions after school over an eight week period during the early and middle part of the second semester. These sessions were held on Monday at School A and on Tuesday at School B. Since participation in the project was voluntary, there was no requirement that teachers attend all the in-service sessions. Therefore, no attendance records were kept on individual teachers. However, the sessions were generally well attended and it was rare for more than 3 or 4 teachers to miss a session.

The in-service training differed in one major respect from the initially planned program. It was originally proposed that the in-service program extend over an eight-week period with two hours per week for a total of 16 hours. It became apparent that this was not going to be possible during the initial discussion of the project with the school principals. Each hour of participation in the program meant an additional hour spent at the school for all the participating teachers. Since the remaining teachers (non-academic subjects) in the two schools were making no comparable investment of time and energy, the problem of "morale" and cooperation among the participants would have been serious for a 16 hour program. Therefore, the 8 hour program was adopted as a compromise between what would be optimal in terms of coverage of topics and what would be optimal in terms of the motivation and cooperation of the participants.

Types of Data Collected

Several different types of data were collected during the course of the project in order to obtain a comprehensive picture of the marking practices of the participating teachers. Each type of data gives unique information about marking and provides answers to questions concerning different facets of the process of assigning marks. The various types of data collected and analyzed in the present project are described briefly in the following sections.

1. Marking Questionnaire Data

A self-report instrument entitled simply "Marking Questionnaire" was constructed and administered to all participants during the first in-service training session. This questionnaire is based upon a previous instrument used in a national survey of the marking practices of approximately 4000 secondary teachers (Terwilliger, 1966). The questionnaire consists of three sections (Part I-Identifying Information, Part II-Marking Practices, and Part III-Attitudes Concerning Marking) and is largely multiple-choice in format. Data from Parts I and II only will be discussed in this report.

In Part I, each teacher was asked to specify the first class he (or she) taught on Monday. _Data for this class only were_ *collected and analyzed in this study*. Therefore, the findings to be reported in following sections are restricted to only one specific class taught by each of the participating teachers. The extent to which the results may have been different if all classes taught by the participating teachers had been included is not known.

The participants were asked to complete the Marking Questionnaire a second time late in the second semester of the 1965-66 school year. A complete copy of the questionnaire is presented in Appendix B.

2. Marking Exercise (Simulation) Data

Two forms (A and B) of a specially designed Marking Exercise were used in the study. The purpose of the exercises was to study the marking behavior of the participating teachers under more controlled conditions than exist in the natural school setting. Each form consists of a description of 20 hypothetical students simulating a class which the participant is to assume he has just taught for a six-weeks period. Brief biographical sketches are provided for each student along with data representing quiz, examination, and daily homework scores. Each participating teacher was asked to assign marks to each of the hypothetical students using the same procedures as with his own students insofar as possible. Form A was administered as a "take home" problem at the beginning of the in-service program and Form B was administered late in the second semester. A copy of Form A and a fuller description of both are given in Appendix C.

3. Standardized Test Data

Historically, standardized aptitude tests came into existence as a means of predicting achievement in the classroom as reflected by marks a student made. Standardized achievement tests came into existence because the marks of different teachers do not necessarily represent the same type and amount of achievement. A vast body of research on the relationship between scores on standardized measures (aptitude and achievement) and average school marks suggests a moderately strong positive relationship. However, little is known concerning the relationship of standardized test scores and the marks assigned by individual teachers. A major question of this study is concerned with the general nature of this relationship and the degree to which it varies from teacher to teacher. Therefore, standardized test data available in the central office at each of the participating schools were collected.

The general procedure was to obtain a complete class roster for the first class taught each Monday by each participating teacher (the only class used for each teacher in this study) and to record from the school files all available standardized test data for each student on each roster. Due to the nature of the testing programs at the two schools, different data are available for students in different grade levels. Table 3 summarizes the types of scores obtained from the records at each school. The only standardized test score available for students at all grade levels in both schools is the Otis Quick Scoring Test of General Ability (IQ).

Table 3

Standardized Test Scores[1] Obtained from Student Files at the Participating Schools

	School A	School B
Grade 9	MAT Word Knowledge Reading Spelling Language Language Std. Skills Arith. Computation Arith. Prob. Solving Soc. Stds. Info. Soc. Stds. Skills Science Otis	MAT Reading Spelling Language Soc. Stds. Info. Arith. Computation Arith. Prob. Solving Science Otis
Grade 10	MAT Reading Spelling Language Soc. Stds. Info. Arith. Computation Arith. Prob. Solving Science Otis	Same as 9th grade
Grade 11	Otis	Same as 9th grade
Grade 12	ACT English Mathematics Social Studies Natural Science Composite Otis	ACT English Mathematics Social Studies Natural Science Composite Otis

[1] Stanine scores were recorded for all Metropolitan Achievement Tests (MAT); ACT standard scores were recorded for all American College (ACT); and standard IQ scores for the Otis Quick-Scoring Test (Otis).

4. Teacher Marking (Grading) Data

The final type of data collected is that which is most central to the project. The actual marks or grades assigned by each teacher to the students enrolled in his (or her) first class each Monday were reproduced from the teacher's grade record book. To assure that all data (including quizzes, homework, tests, etc.) were obtained, actual photographic copies of each teacher's record book were made on copying machines available in the central office at each school. All data available in each teacher's record book for both the first and second semesters of 1965-66 were duplicated. The only exceptions were a few instances in which record books could not be located. These are noted in the discussion of the results.

The amount and relative completeness of the data available in the teacher record books naturally varies greatly from teacher to teacher. In order to analyze systematically the data for all teachers according to a general procedure, it was necessary to limit the amount and type of data included. Inspection of the grade record books revealed that many teachers had recorded a complete set of marks or scores for no more than three assignments or quizzes during a given six-weeks period. Therefore, a decision was made to limit the number of recorded marks or scores to three for assignments and quizzes during any six-weeks period. In addition, a six-weeks daily average, a six-weeks exam, and a final six-weeks mark were recorded for each of the three marking periods in a given semester. Semester exam marks and final semester marks were also recorded. This results in a total of 20 possible marks or scores for each semester. Of course, many teachers had not assigned this many marks.

A brief comment should be made concerning the nature of the data available in the teacher grade record books. With few exceptions, all test, quiz, and other data were recorded by teachers in the form of letter grades (A, B, etc.). This means that no direct numeric scores were available for most teachers. The teachers explained the exclusive use of letter marks resulted from a system-wide policy prohibiting the use of numeric data in any form. Although many teachers (as well as the investigator) expressed disagreement with this policy, few were willing to violate it. Some teachers went to the extent of keeping two separate record books, one for their own use with numeric data and a second for the "official" records with letter grades only. These teachers would allow only the "official" records to be used in this study. This has two important consequences for this study.

First, the exclusive use of letter marks obscures many of the issues which this study set out to clarify. When only letter marks are recorded in a teacher's record book it is not possible to distinguish the measurement process from the evaluation process. That is, concepts such as "difficulty level" and "discriminability" refer to characteristics of classroom tests whereas concepts such as "marking standards" refer to practices employed by the teacher in making value judgments concerning performance on tests and other assignments. Since numeric data are not available in this study, no conclusions can be drawn concerning the measurement devices employed by these teachers. Consequently, it is impossible to determine what role classroom measurement plays in the assignment of marks.

Second, much of the material specifically developed for and used in the in-service program was irrelevant for these teachers. For example, all of the material on the organization and statistical treatment of test and quiz scores is meaningless in a situation where the use of scores is actively discouraged. Similarly, much of the material on test construction and item analysis loses its meaning without the notion of quantifiable test results. Under these circumstances, the potential impact of the in-service program is considerably diminished.

A final (but less important) consequence of the use of letter marks is that all teacher records had to be translated into some arbitrary numeric scale for purposes of data analysis. The scale adopted in this study is described in the next section.

RESULTS

The presentation of the results will be structured according to the various types of data previously mentioned. The data for individual teachers are presented in appendices when they are too complex to be summarized easily in tabular form. Whenever individual data are presented, teachers will be identified by number only. For most types of data, separate summaries for each school and for both schools combined will appear in the text of the report and the more detailed results on which these summaries are based can be found in the appendices. Formal statistical tests have not been used since the primary objective of the study is to explore the nature and magnitude of individual differences among the teachers. The reader who is interested in formal comparisons among various groups of teachers (e.g. School A vs. School B, Language Arts vs. Mathematics, etc.) will usually find the necessary raw data in an appendix.

Standardized Test Data

Perhaps the most meaningful way to contrast the two schools is to examine the standardized test data obtained from the student files. Before doing so, two comments are in order. First, data for only those students in one particular class taught by each participating teacher were obtained from the school files. Second, complete test data were not available for all the students in each class used in the study. For these reasons the data presented here may not be representative of the total student bodies at the two schools.

It was previously noted that different types of standardized test scores were available for the different schools and grade levels within a school (see Table 3). Therefore, the standardized test data are presented separately for each grade level and each school. Tables 4-10 present the sample sizes, means, standard deviations and intercorrelations of standardized test scores for each grade level from 9 through 12 at the two participating schools. The only exception is grade 11 at School A where the only test data available is the Otis. The mean IQ for the 57 students in that group is 118 and the standard deviation is 10.5.

The average Otis IQ is well above that for an unselected sample from the general population in all grades at both schools. As might be expected, there tends to be less variability than in the general population. It appears from inspection of these data that the IQ scores in School A tend to be reliably higher than in School B (especially in grades 10 and 11) despite the fact that both school averages are above the value for the general population. No formal test of the mean difference in IQ was made since these are not random samples from the two schools and because the exploration of such differences was not considered relevant to the present study.

The intercorrelations among the various test scores within each grade tend to be reasonably high although there are perceptable differences between the schools. The median correlation of IQ with all types of achievement scores in School A is .62 and in School B it is .56. The reason for this difference is not known although it may reflect a generally higher (and more uniform)

Table 4

Means, Standard Deviations and Correlations Among Test Scores[1,2]
Ninth Grade--School A (N=89)

Metropolitan Achievement Test

	WK	R	Sp	L	LSS	AC	AP	SSI	SSS	Sci	Otis
WK	---										
R	741	---									
Sp	468	458	---								
L	504	521	576	---							
LSS	429	422	462	741	---						
AC	355	482	522	600	533	---					
AP	344	489	514	627	584	726	---				
SSI	556	694	447	494	470	524	516	---			
SSS	520	553	481	557	423	608	554	682	---		
Sci	509	560	530	349	413	491	459	665	598	---	
Otis	550	557	553	612	531	559	619	560	564	515	---
Mean	7.1	6.9	7.0	6.4	6.3	6.8	6.9	7.0	6.8	6.6	116
SD	1.3	1.7	1.7	1.5	1.4	1.7	1.6	1.5	1.5	1.7	8.3

[1] The following abbreviations are used for subtests of the Metropolitan Achievement Tests:

WK-Word Knowledge
R-Reading
Sp-Spelling
L-Language
LSS-Language Study Skills
AC-Arithmetic Computation
AP-Arithmetic Problem Solving
SSI-Social Studies Information
SSS-Social Studies Study Skills
Sci-Science

[2] Stanine scores were used on all Metropolitan Achievement Tests. Decimal points have been omitted in all correlation matrices.

Table 5

Means, Standard Deviations and Correlations Among Test Scores
Tenth Grade--School A (N=129)

	Metropolitan Achievement Test							
	R	Sp	L	SSI	AC	AP	Sci	Otis
R	---							
Sp	585	---						
L	736	636	---					
SSI	521	395	560	---				
AC	601	540	703	483	---			
AP	623	410	608	550	714	---		
Sci	539	403	527	608	573	554	---	
Otis	699	629	771	655	693	642	628	---
Mean	6.3	6.2	6.0	6.0	5.9	6.0	6.5	112
SD	1.6	1.8	1.9	1.6	2.1	2.0	1.7	11.1

Table 6

Means, Standard Deviations and Correlations Among Test Scores[1,2]
Twelfth Grade--School A (N=77)

	American College Test					
	E	M	SS	NS	C	Otis
E	---					
M	460	---				
SS	527	596	---			
NS	499	635	764	---		
C	735	822	871	877	---	
Otis	613	727	697	707	828	---
Mean	20.9	21.4	23.6	23.2	22.5	117
SD	4.8	5.7	5.6	5.4	4.5	9.3

[1] The following abbreviations are used for the American College Tests (ACT):

E-English
M-Mathematics
SS-Social Studies
NS-Natural Science
C-Composite

[2] ACT standard scores are employed.

Table 7

Means, Standard Deviations and Correlations Among Test Scores
Ninth Grade--School B (N=38)

	Metropolitan Achievement Test							
	R	Sp	L	SSI	AC	AP	Sci	Otis
R	---							
Sp	341	---						
L	419	847	---					
SSI	505	522	446	---				
AC	258	338	250	481	---			
AP	438	539	440	477	862	---		
Sci	469	244	318	517	376	411	---	
Otis	430	498	522	682	555	573	597	---
Mean	7.4	7.4	7.8	7.6	7.2	7.1	6.4	115
SD	1.6	2.0	1.8	1.4	1.9	2.1	1.7	9.5

Table 8

Means, Standard Deviations and Correlations Among Test Scores
Tenth Grade--School B (N=116)

	Metropolitan Achievement Test							
	R	Sp	L	SSI	AC	AP	Sci	Otis
R	---							
Sp	323	---						
L	592	571	---					
SSI	572	225	405	---				
AC	516	364	596	351	---			
AP	396	235	369	255	657	---		
Sci	425	142	358	425	460	489	---	
Otis	494	478	627	327	588	588	534	---
Mean	5.7	5.1	5.1	5.4	4.9	5.1	5.5	105
SD	1.4	1.8	1.5	1.4	1.6	1.6	1.5	10.1

Table 9

Means, Standard Deviations and Correlations Among Test Scores Eleventh Grade--School B (N=87)

Metropolitan Achievement Test

	R	Sp	L	SSI	AC	AP	Sci	Otis
R	---							
Sp	516	---						
L	609	615	---					
SSI	644	374	392	---				
AC	464	384	556	507	---			
AP	639	450	675	575	738	---		
Sci	622	387	468	725	450	598	---	
Otis	655	586	687	522	443	595	552	---
Mean	5.6	5.9	5.9	5.9	6.1	6.2	5.6	107
SD	1.6	1.7	1.5	1.8	1.5	1.4	1.7	10.5

Table 10

Means, Standard Deviations and Correlations Among Test Scores Twelfth Grade--School B (N=38)

American College Test

	E	M	SS	NS	C	Otis
E	---					
M	551	---				
SS	581	311	---			
NS	561	557	642	---		
C	803	806	742	854	---	
Otis	610	646	457	528	710	---
Mean	19.1	19.7	22.0	21.5	20.8	113
SD	4.5	7.6	5.1	5.4	4.6	8.9

achievement motivation in School A which would tend to make general ability a better predictor of achievement. This explanation is also consistent with other known characteristics of the two schools, e.g. the percent of the 1966 graduates who are college bound.

The correlations among the various measures of achievement within each group tend to cluster in the .40-.60 range with occasional higher or lower values. The lowest single correlation among the achievement measures is .22 and the highest (excluding part-whole correlations) is .86. These data certainly suggest that the concept of "general achievement" is a valid one where achievement is defined by performance on measures of this type.

First Semester Grade Data

The data reproduced from the first semester grade record books of the participating teachers were analyzed in several different ways. The first step was to convert all letter marks into numeric values. A 13-point numeric scale was adopted for this purpose since it was deemed desirable to keep distinct all possible letter marks from A+ through F. The following conversion was made for all letter grades: A+=13, A=12, A-=11, ---, D=3, D-=2, and F=1. All analyses of the teacher grading data were made using this 13-point scale.

1. Marks Assigned for Tests, Quizzes, and Other Work During Six-Week Marking Periods

We will begin our consideration of the marking data by restricting attention only to those marks based upon work done within individual six-week marking periods. As previously noted, up to three different quiz, test, or homework marks were recorded for each teacher for each marking period. A daily average and a six-week test mark were also recorded when available for each period. This yields a total of 15 such marks for the three six-week periods in the first semester. The means and standard deviations for all marks recorded by a teacher were computed in all cases where relatively complete data were available. Data from four teachers in School A could not be included in this analysis due to the incompleteness of their records.

Table 11 shows the distribution of these means (using the 13-point scale) by school and type of data, i.e. six-week tests vs. quizzes and daily work. The distributions in Table 11 graphically illustrate the individual differences between teachers (and within a teacher on different occasions) with respect to "standards" of performance. The data in Table 11 constitute marking "norms" based upon a total of 344 quizzes and daily assignments (188 in School A and 156 in School B) and 85 six-week tests (46 in School A and 39 in School B). It is quite clear that teachers in both schools employ higher standards on six-week tests than on daily work and quizzes, i.e. an over-all mean of 6.9 (C+) as compared to 8.0 (B-). There is also evidence that teachers in School A assign higher marks than do those in School B. Perhaps this reflects a conscious attempt to adjust standards of performance to compensate for a higher general ability level at School A. Despite these average differences, there tends to be greater variability within School B than in School A. For example, Teacher 33 (Remedial English) gave 11 quizzes and daily assignments with an

Table 11

Distribution of Mean Class Marks For Quizzes, Daily Work and Six-Week Tests

First Semester

	School A Frequencies		School B Frequencies		Total	
Mean Mark	Daily Work & Quizzes	Six-Week Tests	Daily Work & Quizzes	Six-Week Tests	Daily Work & Quizzes	Six-Week Tests
13.0						
12.5			1		1	
12.0	2		0		2	
11.5	7		6		13	
11.0	10	1	5		15	1
10.5	14	1	5		19	1
10.0	20	3	6	1	26	4
9.5	9	0	7	1	16	1
9.0	16	5	14	3	30	8
8.5	23	2	17	8	40	10
8.0	19	2	20	2	39	4
7.5	17	1	11	4	28	5
7.0	12	8	14	2	26	10
6.5	15	6	11	3	26	9
6.0	10	4	6	3	16	7
5.5	6	6	7	4	13	10
5.0	5	4	6	1	11	5
4.5	1	2	6	1	7	3
4.0	0	1	9	1	9	2
3.5	0		0	2	0	2
3.0	1		3	1	4	1
2.5	0		1	1	1	1
2.0	1		1	0	2	0
1.5				1		1
1.0						
N	188	46	156	39	344	85
Mean	8.4	7.0	7.6	6.7	8.0	6.9
SD	1.8	1.7	2.1	2.1	2.0	1.9

over-all mean grade of 4.2 (D+) while Teacher 48 (Latin I) gave 12 quizzes and daily assignments with a composite mean grade of 10.6 (A-). Similar differences obtain for six-week tests when specific extreme teachers are selected.

A second type of analysis was performed for a subsample of five teachers from each school. Multiple regression equations predicting six-week marks from all data recorded in each marking period were determined. This yields two important pieces of information; (a) the degree of predictability of assigned six-week marks, and (b) the relative weight each piece of data carries in the prediction of the assigned mark. Since this analysis was exploratory and involved only ten teachers, the results and discussion are presented in Appendix D.

2. Marks Reported to Parents

The most important marks to the student are those which appear on his report card. The marks reported each semester to students and their parents in the secondary schools of Nashville are the three six week marks, the semester examination mark, and the final semester mark. The means and standard deviations for each of these marks are given for each teacher in Table 12. The only exceptions are Teachers 3, 21, 41, and 42. In each of these cases the appropriate data were not available when the record books were duplicated. These data reveal in great detail the nature and extent of differences among individual teachers with respect to their practices in assigning marks. For example, contrast the mean marks assigned by Teacher 17 with those assigned by Teacher 14. The mean for the first teacher is about 9.0 (B) and for the latter it is 5.5 (between C- and C). This is a substantial difference and is not entirely explained by differences in the ability levels of these classes. (The mean IQ in the first class is 121 and in the second it is 113.)

Table 12

Means and Standard Deviations of Reported First Semester Marks

I.D. No.	N	1st 6 Wks		2nd 6 Wks		3rd 6 Wks		Sem. Exam		Sem. Mark		Subject
		M	SD	M	SD	M	SD	M	SD	M	SD	
01	27	5.3	2.7	5.9	2.6	6.2	2.9	5.6	3.9	5.8	3.0	Biology
02	30	8.3	2.7	7.8	2.4	8.6	3.3	5.8	3.2	8.0	2.7	Am. Govt.
04	27	7.8	3.0	8.6	2.3	8.1	3.0	7.7	3.6	7.9	2.9	Spanish
05	16	7.6	2.2	7.4	2.0	8.0	3.3	7.6	2.6	7.7	2.2	French
06	17	8.7	1.4	8.2	1.8	7.8	1.6	7.6	2.5	8.2	1.7	French
07	23	10.0	.9	10.3	.9	9.0	1.8	8.6	1.8	9.6	1.2	English
08	27	8.2	2.1	8.0	2.4	7.5	1.5	9.0	2.8	8.3	2.0	English
09	25	6.8	3.6	6.6	3.5	5.9	3.4	5.0	4.4	5.9	3.5	History
10	28	7.3	3.1	7.3	2.9	7.4	3.1	6.0	3.4	7.2	3.1	Spanish
11	21	7.2	2.5	7.5	3.0	7.7	2.4	5.0	3.1	6.8	2.7	Biology
12	27	6.8	2.7	7.6	2.6	8.2	2.7	8.4	2.8	8.1	2.7	Latin
13	26	8.2	2.1	7.7	3.3	5.3	3.2	7.4	3.5	7.2	2.8	Algebra
14	24	6.2	2.4	5.3	2.4	6.2	2.3	5.0	2.6	5.7	2.2	English
15	24	10.7	2.2	10.4	2.1	10.0	2.7	8.1	2.7	9.8	2.5	Algebra
16	31	7.2	2.9	7.4	2.5	7.8	2.4	7.1	3.3	7.5	2.4	Sociology
17	30	10.3	1.4	10.3	1.9	10.3	1.5	7.1	2.8	8.8	2.0	History
18	18	9.0	3.4	6.6	4.0	7.5	2.9	5.8	2.9	7.3	3.0	Bus. Arith.
19	28	8.2	1.9	7.1	2.4	8.8	2.2	5.1	3.6	7.4	2.2	Algebra
20	23	9.2	2.2	7.7	2.4	9.2	2.5	7.2	3.1	7.9	2.4	[1] Biol.(Special)
30	20	7.8	3.0	7.7	3.0	8.0	2.9	6.2	4.7	7.2	3.2	Adv. Math.
31	32	6.5	3.2	5.2	3.2	4.9	3.4	5.2	2.7	5.3	2.9	Geometry
32	27	5.0	3.0	5.5	2.7	5.7	2.8	5.2	3.4	5.4	2.7	Biology
33	23	2.9	1.7	4.8	1.5	3.7	1.9	4.4	2.0	4.4	1.7	[2] Eng.(Special)
34	25	6.8	3.5	7.6	3.6	7.5	3.7	5.8	4.3	7.2	3.6	English
35	27	8.2	3.0	6.7	2.9	6.0	2.3	5.0	4.0	6.2	2.5	Biology
36	32	6.1	2.4	5.7	2.8	5.2	2.7	5.9	3.1	5.8	2.4	Geography
37	23	8.6	3.0	8.9	2.2	8.2	2.7	8.4	3.0	8.5	2.6	Am. Govt.
38	27	7.0	3.7	6.6	3.4	7.7	2.7	6.3	3.7	6.7	3.0	History
39	32	7.2	2.8	6.9	2.9	7.1	2.7	8.0	4.1	7.1	2.8	History
40	26	8.5	2.6	7.4	3.4	6.6	3.2	7.6	3.8	7.4	3.2	Algebra
43	14	7.5	2.4	8.6	1.7	8.4	1.6	7.6	2.6	7.9	1.9	Geography
45	24	8.2	3.4	8.9	2.5	8.3	2.5	6.5	3.7	7.7	3.1	Spanish
46	29	6.1	3.1	4.7	2.9	10.3	1.4	3.3	2.6	5.7	3.1	English
47	26	8.3	3.0	8.9	3.7	7.4	3.8	6.6	3.2	8.0	3.2	History
48	26	9.5	2.9	8.8	2.6	8.8	2.7	4.9	4.3	8.0	2.8	Latin

[1] Low ability group

[2] Low ability remedial section

A more compact summary of the data in Table 12 is provided in Table 13 where the grand mean grade and the variability among the mean grades is given for each six-week period, the semester exam, and the final semester mark. Table 13 again reveals that teachers in School A tend to mark higher than do teachers in School B. The average in School B shows a remarkable consistency across the three grading periods but, as with the quiz and daily marks, there tends to be slightly greater variation from teacher to teacher than in School A. In both schools there is a noticeable drop in semester exam marks as compared with other reported marks.

Despite the overall high ability level, there remained considerable variation in mean IQ from class to class.

A check was made to see if the variability in the average final mark might be explained by differences in the average general ability of the classes. The correlation between average class IQ and average final mark was computed for both schools combined (N=35) and found to be .57. Therefore, slightly over one-fourth of the variability in average assigned marks can be accounted for by differences in the average ability as measured by the Otis. A substantial portion of the variability in average marks is evidently due to other factors.

The intercorrelations of IQ and marks for each six weeks, final exam and total semester were also obtained within each class with at least 15 students with complete data. Sixteen teachers in School A and 12 in School B met this criterion. The resulting matrices are shown in Appendix E. The correlations between IQ and final semester marks are summarized in Table 14.

Table 13

Means and Standard Deviations of Average Reported Marks First Semester

		1st 6 Wks	2nd 6 Wks	3rd 6 Wks	Sem. Exam	Sem. Mark
School A (N=19)	M	8.1	7.8	7.9	6.8	7.6
	SD	1.4	1.3	1.3	1.3	1.1
School B (N=16)	M	7.1	7.1	7.1	6.1	6.8
	SD	1.6	1.5	1.6	1.3	1.1
Total	M	7.6	7.5	7.5	6.5	7.2
	SD	1.5	1.5	1.5	1.4	1.2

Table 14

Distribution of Correlations of Otis IQ with First Semester Marks

r	School A	School B	Total
.75			
.70		2	2
.65	3	3	6
.60	1		1
.55	1		1
.50	1	1	2
.45	2		2
.40	1		1
.35		2	2
.30		2	2
.25	4	2	6
.20			
.15			
.05	1		1
.00			
-.05	1		1
-.10	1		1
N	16	12	28
M	.36	.47	.41
SD	.24	.18	.22

It is difficult to generalize from samples as small as those given in Table 14. However, two features of the data deserve comment. The average correlation of IQ and final marks is substantially higher and the variability in this relationship is less from teacher to teacher in School B. The distributions tend to be bimodal in both schools with one group of teachers clustered in the .60-.70 range and a second group clustered around .25-.35.

Since the general ability level is higher in School A, it seems reasonable to assume that the lower average correlation of IQ and marks might be a result of attenuation due to restriction of range in ability. Specifically, those cases in which the correlation is very low may simply be the most able (and least variable) classes with respect to the Otis. This is borne out in the two cases where small negative correlations were obtained. In each instance, the mean IQ is greater than 120 and the standard deviation is about 6.0. Yet, some caution is needed in offering this explanation for all low correlations because counter-examples are also available, e.g. the correlation between IQ and marks is .40 for one class with a mean IQ of 124 and a standard deviation of 5.0 and is .60 for a second class with a mean of 121 and a standard deviation of 9.0.

It is also instructive to contrast the magnitude of the Otis-mark correlations with the intercorrelations among the marks for individual teachers (see Appendix E). Although no attempt has been made to formally summarize these, it is obvious that the Otis is not accounting for a large part of the general variance in marks since the intercorrelations among marks are typically in the .80-.90 range while the correlations with the Otis are much lower. This tends to be true regardless of mean IQ level in a class.

A final analysis of the marking data consists of deriving a multiple regression equation for each teacher predicting semester marks from marks for the three six-week periods and the semester exam. The zero-order correlations among these variables are given for each teacher in Appendix E. The purpose of this analysis is to determine the over-all predictability of semester marks from other previously recorded data and to ascertain the relative weight given to each type of data in the determination of semester marks. Multiple regression equations were obtained for 19 teachers from School A and 16 teachers from School B. The remaining teachers did not have complete data for all the appropriate variables.

The results of the multiple regression analyses are presented in Table 15. The values in the right-hand column indicate that semester marks are highly predictable from previous marks. This is not surprising since teachers typically average previous marks in some fashion to obtain a final mark. Of course, if the final marks actually represent some linear combination of the six-week and final exam marks then they would be perfectly predictable. This is approximately the case for most teachers. In specific instances where a substantial proportion of the variance in final marks is not accounted for (e.g. Teacher 33) it may be that other factors such as special projects or term papers entered into the determination of marks.

Table 15

Percent of Variance in First Semester Marks Accounted for by Period and Semester Exam Marks

Teacher Number	Six-Week Period			Semester Exam	Total
	1	2	3		
1	12	22	33	32	99
2	21	9	35	28	93
4	29	10	28	30	97
5	20	17	33	29	99
6	5	40	15	37	97
7	15	14	36	28	93
8	37	10	11	38	96
9	17	34	28	19	98
10	21	15	28	28	92
11	5	24	18	46	93
12	1	7	87	4	99
13	6	29	32	23	90
14	25	31	25	17	98
15	10	24	40	24	98
16	0	17	26	51	94
17	2	11	25	56	94
18	14	22	38	23	97
19	21	25	23	28	97
20	22	37	20	21	100
30	15	11	23	48	97
31	12	25	36	22	95
32	26	28	14	28	96
33	22	34	-5	31	82
34	5	27	50	17	99
35	11	24	24	39	98
36	9	13	33	38	93
37	24	11	27	36	98
38	35	21	21	23	100
39	0	15	66	15	96
40	7	29	36	27	99
43	31	30	-13	50	98
45	12	20	20	43	95
46	56	13	6	16	91
47	23	22	23	28	96
48	19	22	30	27	98
Mean (A)	15	21	30	30	96
Mean (B)	19	22	24	31	96
Mean (Total)	17	21	28	30	96

The general ordering of the four predictors with respect to "variance accounted for" is the same in both schools. The first two marking periods are relatively less important than the third period and the final exam. There is a slight tendency for the total variance to be more evenly distributed across the four predictors in School B.

The data also clearly indicate that there are extreme individual deviations from the general trends just noted. Teacher 12, for example, apparently looked almost exclusively at work done in the third six-week period since 87% of the final marks are accounted for by that period. The same is true to a lesser extent for Teacher 39. Teachers 33 and 43 represent rare examples where one predictor (third six-week mark) actually contributes negatively to the overall prediction. This means that this predictor is partialling out variance in final marks which the other predictors have previously accounted for. The net result is that the total variance accounted for is reduced.

Marking Exercise (Form A)

All participating teachers except one completed the Marking Exercise (Form A) in the week between the first and second inservice sessions. The primary purpose in collecting these data was to obtain marks assigned by each teacher under known and standardized conditions. This, of course, is not possible in the natural classroom setting where each teacher must deal with a unique group of students using his own measurement and evaluation procedures. The distributions of assigned marks for the 20 hypothetical students in Form A of the Marking Exercise are shown in Table 16. These data are for both schools combined since the difference in results between the schools was slight.

Table 16

Distribution of Assigned
Marks for Marking Exercise (Form A)

Student Number	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A	A+	M	SD
1						7	8	8	14	1				7.8	1.2
2									2	3	11	22		11.4	.8
3									8	7	8	15		10.8	1.2
4						21	7	5	5					6.8	1.1
5						3	2	5	24	3		1		8.7	1.1
6						6	4	6	20	1	1			8.2	1.2
7			1			19	9	4	5					6.8	1.2
8			1			16	6	7	8					7.1	1.4
9			3	2	5	17	6	3	2					6.0	1.4
10									9	8	9	12		10.6	1.2
11			13	3	9	13								4.6	1.3
12	5	3	14	2	8	6								3.6	1.6
13		1	11	4	7	15								4.6	1.3
14			7	4	11	14	2							5.0	1.2
15									6	9	5	18		10.9	1.2
16			4		5	21	5	1	2					5.9	1.3
17			8	2	11	16		1						5.0	1.2
18			10	6	7	15								4.7	1.2
19			6	1	8	16	3	1	3					5.6	1.6
20								1	6	7	7	17		10.9	1.2

Inspection of the distributions in Table 16 shows obvious and consistent discriminations being made among the 20 students. This is also reflected in the great variation in the mean marks for the 20 students and the relatively low variability of the marks for most students. The consistency among the marks of the 38 teachers is also demonstrated in the intraclass correlation (R) which is .80 for these data. (R was computed so that the average differences among teachers were treated as error.) Viewed in the context of what is known about rating techniques, these data would appear to reveal a relatively high degree of agreement among teachers.

However, from another point of view, these data are not necessarily encouraging. It could be argued that an R=.80 is not really so high when one considers the fact that all teachers were given the identical data on which to base their marks. Under these circumstances, the variability in the marks given to certain students (e.g. Students 9, 12, and 19) seems unreasonable. The explanation of such variability undoubtedly lies in differences among teachers in the subjective values which are attached to different types of data available for each student.

Multiple regression equations were computed for each teacher in order to determine the effective weight given to each of the seven scores available from the Marking Exercise. Table 17 gives the percent of variance in assigned marks accounted for by each of the scores given in the exercise. There is one important difference between these data and the data given in Table 15. Whereas each of the predictors and the criterion variable were generated by the individual teacher in the analysis of the first semester marks, only the criterion variable is generated by the teacher in this analysis. The means, standard deviations, and correlations among the predictors are fixed and are the same for all teachers in the present case (see Appendix C). To a certain extent the range of weights which these predictors can carry is restricted. However, a teacher who combines the predictors in specific ways can also exercise a great deal of control over the weights. One basic purpose of this analysis was to determine the extent to which individual teachers succeeded in producing weights which reflect the importance they attach to various types of data.

Table 17

Percent of Variance in Assigned Marks Accounted for by the Seven Scores in Marking Exercise A

I.D.	Quiz Number				3 Wk	6 Wk	Homework	Total
No.	1	2	3	4	Test	Test	Average	
1	11	12	25	25	21	7	-12	89
2	11	- 5	20	43	13	10	- 6	86
3	19	12	0	5	31	21	11	99
4	- 1	- 6	6	8	11	57	19	94
5	7	12	16	2	22	35	0	94
6	7	13	11	4	13	41	18	99
7	1	5	11	31	14	28	9	99
8	1	23	1	3	7	24	25	82
9	22	7	7	20	21	22	- 1	98
10	- 3	13	10	9	13	29	18	89
11	20	- 8	7	25	10	15	27	96
12	29	4	11	19	8	24	- 6	89
13	12	11	5	23	3	4	40	98
14	9	3	-23	33	21	31	20	94
15	6	-11	4	56	0	- 6	40	89
16	4	14	3	18	15	33	7	94
17	3	5	8	24	30	14	13	97
18	23	16	9	20	10	17	1	96
19	17	13	18	30	6	14	1	99
20	17	8	0	21	19	21	13	99
21	14	15	7	- 5	46	28	- 8	97
30	1	5	11	31	14	28	9	99
31	27	10	0	26	14	17	3	97
32	9	7	2	28	14	21	17	98
33	- 1	16	- 4	49	- 6	18	24	100
34	4	9	- 2	23	11	26	27	98
35	30	23	4	28	5	- 7	11	94
36	8	6	7	14	33	24	4	96
37	11	5	26	33	12	10	2	99
38	5	28	9	30	4	9	5	90
39	20	16	21	- 1	30	13	- 7	92
40	12	9	2	26	13	25	10	97
41	13	1	-18	47	25	12	15	95
42	9	4	21	24	3	26	4	91
45	26	16	3	2	27	19	- 1	92
46	- 2	5	- 4	14	43	34	8	98
47	5	9	2	9	10	31	27	93
48	18	7	- 8	30	16	19	13	95
Mean	11.2	8.7	6.0	21.8	15.8	20.9	10.5	

The values in the right-hand column of Table 17 indicate that most teachers relied almost exclusively upon the seven scores for each student in assigning marks. Only a few teachers (1, 2, 8, 10, 12, and 15) assigned marks based upon other factors in addition to these scores. This implies that the biographical data given for each student had little impact in the marking.

The column means at the bottom of the table indicate that the scores on the last quiz and the six-week test each accounted for 20 percent of the variance over all teachers. The 3-week test, the first quiz, and the homework average are next in order of importance and the second and third quiz are least important. These data illustrate the importance of the variability and intercorrelations among scores which enter into a composite. Due to the way in which the exercise was constructed, the last quiz has the greatest variability of all predictors and also has a high average correlation with the remaining predictors. Therefore, it is almost certain to have a heavy impact on any type of average both directly (because of its great variability) and indirectly (through its correlation with the remaining predictors). This effect will tend to manifest itself despite the fact that extra nominal weights are typically assigned to the more major types of test data. Of course, the variances and average correlations of the other predictors play similar roles in determining the weights they have in the multiple regression. The only thing that a teacher can do to control these influences is to convert all scores to a standard score scale (thereby equating variances) and give the converted scores the nominal weights which are judged appropriate. The data in Table 17 clearly indicate that teachers do not generally use such conversion techniques. As a consequence, the last quiz accounts for a disproportionate share of the variance in assigned marks, e.g. Teachers 2, 15, 33, and 41. Conversely, the major six-week test often accounts for little of the variance because of failures on the part of teachers to use weighted converted scores, e.g. Teachers 1, 13, 15, 35, and 38. The obvious conclusion is that these teachers have little insight into the significance of differences in the variability of scores with which they were asked to work.

The Marking Exercise data were also used to check the possibility that the marking standards employed by teachers represent generalized habits which are consistent across a variety of situations. If such is the case, the average marks assigned on the Marking Exercise should correlate with the average marks assigned to actual students in class. The correlation between mean first semester mark and mean mark assigned to the hypothetical students was found to be .05. Therefore, marking standards do not generalize over situations such as these.

Marking Questionnaire Data (Time 1)

The frequency of responses to each item of the Marking Questionnaire were converted into proportions and these proportions are given in Appendix . These proportions are based upon 41 respondents including the 39 participants, Teacher 44, and a practice teacher at School B. Since these proportions are self-explanatory no extensive comment will be made. However, for purposes of comparison, proportions based upon a previous survey of approximately 2500 secondary teachers are also provided. In both the original survey and the present study great variation was found in the responses of teachers to specific questions. These data suggest the need for greater standardization of practices in the whole marking process. As previously noted, the data for Part III of the questionnaire are not given because no norms are available for comparison.

The remaining portion of the presentation of results is devoted to data collected after the in-service training sessions had concluded. The first six-week marks of the second semester were assigned by teachers during the in-service training and the second six-week marks were assigned shortly after the conclusion of the in-service program. Therefore, if the program had any impact upon marking practices its influence should be obvious in most marks assigned during the second semester. Also, the effects of the program should be manifested in the responses to Marking Exercise B and the repeat administration of the Marking Questionnaire late in the second semester.

However, it should be noted that differences which are found between first and second semester data cannot automatically be attributed to the in-service program because no control group was used. It is possible that there are generalized differences in the marking practices of teachers during the two semesters of a course. Only tentative statements are possible concerning the impact of the program.

Second Semester Grade Data

1. Marks Assigned for Tests, Quizzes, and Other Work During Six-Week Marking Periods

Table 18 gives the distribution of average marks assigned by teachers on daily work, quizzes, and six-week exams during the second semester. These data are based upon the records of 31 teachers for whom relatively complete data are available. The grand mean mark is consistently lower for both schools and both types of data when compared with the first semester data in Table 11. The variability in six-week test marks is greater in School A than School B whereas the opposite was true in the first semester. These differences may be due to the fact that the data for the two semesters are not based upon identical groups of teachers since the completeness and availability of the record books varied from one semester to the next. Under these circumstances, any statements concerning the impact of the in-service program would be extremely hazardous.

The special multiple regression analysis on the data within each six-week period was repeated for the same subsample of teachers used with the first semester data. The results and discussion for both semesters are in Appendix D.

2. Marks Reported to Parents

The means and standard deviations of the second semester marks assigned by each teacher during each period, on the final exam, and for the total semester are shown in Table 19. Data for five teachers (2, 13, 16, 21, and 36) were not available. The corresponding individual data for the first semester are in Table 12. No attempt will be made to make comparisons between the two sets of data for individual teachers.

Table 18

Distribution of Mean Class Marks For Quizzes, Daily Work and Six-Week Tests

Second Semester

Mean Mark	School A Frequencies Daily Work & Quizzes	School A Frequencies Six-Week Tests	School B Frequencies Daily Work & Quizzes	School B Frequencies Six-Week Tests	Total Daily Work & Quizzes	Total Six-Week Tests
13.0						
12.5			3		3	
12.0	3		2		5	
11.5	3	2	5		8	2
11.0	9	0	3		12	0
10.5	10	0	6		16	0
10.0	18	1	3		21	1
9.5	17	3	3	2	20	5
9.0	20	2	10	2	30	4
8.5	15	4	20	4	35	8
8.0	13	1	14	0	27	1
7.5	12	3	14	2	26	5
7.0	17	4	12	5	29	9
6.5	13	1	5	4	18	5
6.0	17	7	8	2	25	9
5.5	9	3	9	5	18	8
5.0	6	2	5	3	11	5
4.5	5	2	4	2	9	4
4.0	0	1	5	5	5	6
3.5	2	2	4	1	6	3
3.0	1	0	2		3	0
2.5		1	4		4	1
2.0			1		1	
1.5						
1.0						
N	190	39	143	37	333	76
Mean	8.1	6.9	7.5	6.3	7.8	6.6
SD	1.8	2.1	2.4	1.7	2.2	2.0

Table 19

Means and Standard Deviations of Reported Second Semester Marks

I.D. No.	N	1st 6 Wks M	1st 6 Wks SD	2nd 6 Wks M	2nd 6 Wks SD	3rd 6 Wks M	3rd 6 Wks SD	Sem. Exam M	Sem. Exam SD	Sem. Mark M	Sem. Mark SD
01	25	5.4	2.6	6.3	2.6	6.4	2.5	7.0	4.3	6.4	2.9
03	25	7.8	2.2	7.3	1.5	7.7	2.4	7.8	4.0	7.8	2.2
04	27	8.6	2.3	7.8	2.7	6.7	2.7	4.4	3.5	6.8	2.7
05	16	6.5	2.5	7.2	2.4	7.1	2.2	5.6	3.1	6.8	2.3
06	17	8.9	1.6	7.1	2.0	8.1	1.6	6.9	3.2	7.7	2.0
07	24	10.1	1.0	10.2	1.0	9.6	1.6	9.4	1.5	9.8	1.0
08	28	6.1	3.6	8.5	2.6	8.3	2.5	7.3	3.7	7.9	3.1
09	25	6.1	3.1	4.5	3.4	4.8	3.5	4.1	4.2	5.0	3.5
10	27	7.8	3.4	6.8	3.0	7.4	3.3	4.2	3.1	6.6	3.1
11	20	9.0	2.2	8.2	2.8	9.4	2.0	7.1	3.6	8.0	2.4
12	26	8.0	2.3	7.8	2.7	8.4	2.3	8.4	3.1	8.4	2.6
14	23	5.7	2.2	5.5	2.1	7.3	2.1	4.8	2.7	6.0	2.2
15	26	8.1	3.3	8.6	3.0	7.9	2.8	9.6	3.0	8.6	3.1
17	28	10.5	1.8	10.9	1.4	10.1	2.6	9.6	2.5	9.9	1.8
18	18	6.9	3.8	6.8	3.1	6.4	3.4	6.4	4.1	6.2	3.5
19	28	8.1	2.8	7.8	2.8	9.0	1.7	5.8	3.2	7.9	2.2
20	26	8.2	3.1	7.4	2.5	8.7	2.5	7.0	3.1	7.8	2.5
30	33	5.9	3.7	5.4	3.6	6.0	3.5	5.0	3.9	5.5	3.5
31	18	4.4	3.5	4.2	3.5	5.1	3.4	4.7	3.5	4.7	3.6
32	29	7.7	2.5	5.3	2.6	7.3	2.9	5.0	3.4	6.2	2.5
33	21	3.5	1.5	2.8	1.9	3.8	2.0	3.5	2.1	3.4	1.6
34	24	5.8	4.6	6.9	3.7	6.9	3.6	5.5	4.2	6.4	3.9
35	29	8.7	4.5	7.9	2.8	7.8	2.8	7.2	3.7	7.3	3.4
37	22	8.4	2.5	8.4	1.9	7.9	2.5	9.0	2.5	8.6	2.2
38	29	7.0	4.0	4.6	3.4	6.1	3.2	6.7	3.9	6.1	3.5
39	32	5.9	3.4	6.3	3.3	6.4	3.2	6.0	4.4	6.0	3.6
40	26	6.2	3.4	6.3	3.4	5.9	3.2	6.6	3.5	6.4	3.2
41	24	6.9	3.5	6.5	2.9	8.0	2.4	4.9	3.8	7.1	2.7
42	30	6.1	2.8	2.6	1.7	5.6	3.0	7.1	3.3	5.6	2.3
43	17	6.4	2.4	5.6	1.9	6.2	3.0	4.9	3.1	5.8	2.8
45	24	8.7	2.9	8.6	2.9	8.1	3.5	7.0	4.5	8.2	3.6
46	26	6.6	3.0	7.1	2.2	6.6	2.2	4.2	3.4	5.8	2.8
47	25	9.1	3.6	7.6	3.7	8.0	3.7	7.2	3.6	8.0	3.2
48	25	7.6	2.7	7.4	3.1	9.2	2.4	5.0	4.3	7.5	3.1

The summary data for each school and the total sample are presented in Table 20. The comparable first semester summary is in Table 13. As in the first semester, the average marks assigned at School B are consistently lower than at School A, and semester exam marks are lower than other marks at both schools. The second semester averages at both schools are consistently lower than the corresponding first semester values. This reflects the same general finding previously noted with respect to data on quizzes, daily work, and six-week exams. A check was made to determine whether this difference could be accounted for by the fact that certain teachers contributed data for only one semester. The means for marks presented in Tables 13 and 20 were recomputed using data from only the 16 teachers in School A and the 15 teachers in School B who had data both semesters. The results are nearly identical with those in Tables 13 and 20. The decrease in average marks in the second semester represents actual changes in the standards of the teachers rather than differences in the samples used in the two semesters. It is not possible to state whether this is a general phenomenon or is a result of the in-service training. The change is, however, consistent with an emphasis in the program on the need for an adjustment of the average difficulty level of tests and assignments to the average ability in the group. For most teachers, this implies making such tasks more difficult.

The average final marks for the second semester were correlated with average class IQ for all teachers where both were available (N=32). This correlation is .66 as compared with the .57

obtained for the first semester marks. This represents an increase of 11 percent in "variance accounted for ". The individual differences in average second semester marks assigned by teachers tend to be more closely related to the average ability level of their students than do the individual differences in first semester average marks. Again, the in-service training program cannot be tied to this difference in a causal manner but the nature of the change is consistent with one of the primary objectives of the program, viz. to make teacher's standards more sensitive to the ability levels of the groups they teach. It should be noted that the increase in relationship between average marks and average IQ did not result in an extensive change in the individual differences in average marks since the correlation between average marks for the two semesters (N=31) is .82.

Table 20

Means and Standard Deviations of
Average Reported Marks
Second Semester

		1st 6 Wks	2nd 6 Wks	3rd 6 Wks	Sem. Exam	Sem. Mark
School A (N=17)	M	7.8	7.6	7.8	6.8	7.5
	SD	1.4	1.5	1.3	1.8	1.3
School B (N=17)	M	6.8	6.1	6.8	5.8	6.4
	SD	1.5	1.7	1.3	1.4	1.3
Total (N=34)	M	7.3	6.8	7.3	6.3	7.0
	SD	1.5	1.8	1.4	1.6	1.4

The correlations among the reported second semester marks and IQ scores were obtained for all teachers with at least 15 students having complete data. The correlation matrices for the

24 teachers meeting this criterion are given in Appendix E. The correlations between IQ and final second semester mark have been abstracted from these matrices and are summarized in Table 21. The corresponding summary for the first semester data is in Table 14. The average correlation in School B is higher but the variability from teacher to teacher is less. This is consistent with the first semester data. However, comparing across semesters, we find the average correlation in both schools slightly lower for the second semester. The variability in School A increased and the variability in School B decreased slightly during the second semester. The change in the average correlation for both schools combined is so small (.03) that, for practical purposes, the average correlations for the first and second semesters may be considered equal. It is apparent from these data that, regardless of changes on the part of individual teachers, the over-all picture is one of no change. The impact of the in-service program on the average IQ-final mark correlation is therefore negligable.

The multiple regression analysis performed on the first semester final marks was repeated with second semester marks. Regression equations were derived for each teacher in the prediction of final marks from the three six-week period marks and the semester exam mark. The results are given in Table 22 in terms of the final mark variance accounted for by each of the predictors. The averages for the combined groups indicate that the semester marks account for almost one-third of the variance while the third six-week marks account for about 28 percent. The first and second period marks each account for slightly less than 20 percent of the

Table 21

Distribution of Correlations of Otis IQ with Second Semester Marks

r	School A	School B	Total
.90	1		1
.85			
.80			
.75	1		1
.70			
.65	1	1	2
.60		1	1
.55	1	1	2
.50		2	2
.45		1	1
.40	2	1	3
.35		1	1
.30	1	2	3
.25	1		1
.20			
.15	1	1	2
.10	1		1
.05	1		1
.00			
-.05	1		1
-.10			
-.15	1		1
N	13	11	24
M	.33	.43	.38
SD	.30	.14	.25

Table 22

Percent of Variance in Second Semester Marks Accounted for by Period and Semester Exam Marks

Teacher Number	Six-Week Period 1	2	3	Semester Exam	Total
1	29	4	26	37	96
3	8	18	25	46	97
4	16	16	30	34	96
5	7	28	25	35	95
6	17	23	13	46	99
7	14	4	44	32	94
8	12	1	40	42	95
9	9	27	38	25	99
10	19	22	19	36	96
11	5	- 5	44	53	97
12	21	10	18	48	97
14	11	- 8	51	42	96
15	7	45	27	19	98
17	12	14	30	32	88
18	21	3	45	28	97
19	29	29	15	25	98
20	26	17	21	33	97
30	34	28	14	24	100
31	21	31	17	30	99
32	18	26	20	34	98
33	10	38	20	19	87
34	18	14	32	35	99
35	36	- 7	35	25	89
37	15	22	25	35	97
38	27	24	14	32	97
39	13	7	27	51	98
40	23	30	20	26	99
41	30	11	36	17	94
42	34	0	22	33	89
43	6	23	39	31	99
45	4	10	62	23	99
46	22	15	12	41	90
47	8	22	20	41	97
48	33	45	22	- 2	98
Mean (A)	16	15	30	36	97
Mean (B)	21	20	26	30	97
Mean (Total)	18	17	28	33	96

variance. This outcome is very similar to the results for the first semester (see Table 15), the only difference being that slightly more variance is being predicted by the final exam and slightly less is being predicted by the second period marks. The average total variance accounted for is the same in both semesters.

These statements refer to average findings and, of course, changes at the individual teacher level tend to be obscured by such averages. Significant changes by individuals (e.g. Teachers 12, 33, and 48) did occur but it is difficult to ascertain whether such changes reflect radically different practices consciously adopted by the teachers or "chance" differences in the pattern of correlations among the six-week period and final exam marks.

Marking Exercise (Form B)

The second Marking Exercise was administered during the last three weeks of the second semester. Since the in-service program had concluded, the materials were left in each teacher's mailbox with a request to complete and return the exercises prior to the end of the semester. This request unfortunately came at a time when most teachers were under great time pressure due to special end-of-the-year activities and events, final examinations, etc. Despite two follow-up requests during the final two weeks of the term only 23 (15 from School A and 8 from School B) completed copies of the Marking Exercise were returned. This, in combination with differences inherent in the two forms, make it virtually impossible to make direct comparisons between the data collected on the two occasions. The distributions of assigned marks for Form B are presented in Table 23. As with Form A, these data clearly show that fairly consistent discriminations are being made between students. However, the Form B intraclass correlation of .69 is appreciably lower than the .80 obtained for Form A. This difference may reflect sampling error in the two groups completing the two forms or it may be attributed to differences in the two forms themselves (see Appendix C).

Table 23

Distribution of Assigned
Marks for Marking Exercise (Form B)

Student Number	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A	A+	M	SD
1			1	1		15	2	1	3					6.4	1.4
2								2	16	1	1	3		9.4	1.1
3						1		2	15	1	2	2		9.3	1.3
4			2	1	2	16	1	1						5.7	1.1
5						11	5	6	1					6.9	.9
6			1		1	14	3	1	3					6.4	1.3
7			5		5	12	1							5.2	1.2
8		1	2	2	1	15		1	1					5.6	1.5
9			8	3	4	7			1					4.6	1.5
10						1	1	2	16	2	1			8.9	.9
11	1		16		2	4								3.6	1.3
12	3	5	11	2	1	1								2.8	1.2
13	2	2	15		3	1								3.1	1.2
14	1		11	2	3	6								4.0	1.4
15							2		19	1	1			9.0	.7
16			9	2	3	8	1							4.6	1.4
17	2		15	1	1	4								3.5	1.4
18	2	1	12	1	2	4			1					3.8	1.8
19			4	3	8	7			1					5.0	1.4
20							1	2	14	1	2	3		9.4	1.3

The fact that Form B is "harder" (in the sense that scores are systematically reduced on five of the seven classroom performance variables) is clearly reflected in the mean marks assigned on Form B. The average mark for every one of the 20 students is less than the corresponding average for Form A. The grand mean for all marks on Form A is 7.3 as compared with 5.9 for Form B. This finding has one obvious implication -- the teachers gave an absolute rather than normative interpretation of the magnitude of the classroom performance scores. That is, the 0-100 score scale used in the exercises was treated as if it represents meaningful units which could be interpreted without reference to the performance of the group. The absolute interpretation undoubtedly represents a carry-over from the system where 93-100=A, 85-92=B, etc. The fallacy of absolute interpretation of scores was discussed in the in-service program but this apparantly had little impact upon the teacher's actual behavior. Perhaps this can be partially explained by the general absence of numeric data when the teachers assign marks to their own students. Where numeric data are not employed, it is difficult for teachers to see the need for using basic descriptive statistics in order to interpret test and quiz results. The long-term habits established under these conditions were evidently unaffected by the in-service training.

The multiple regression analysis was repeated on the Form B data for each teacher. Table 24 shows the percent of variance in assigned marks attributable to each of the seven classroom scores. The average total percent of variance accounted for is 90% as opposed to 95% for Form A. The general reduction in the scores may have forced some teachers to rely more upon biographical data to avoid assignment of marks which they consider unduly low. Had the samples been identical for the two forms and had each teacher employed the same method on both occasions, we would expect Table 24 to be a rearrangement of the columns in Table 17 corresponding to the shifting of variables in Form A to construct Form B.

Table 24

Percent of Variance in Assigned
Marks Accounted for by the Seven
Scores in Marking Exercise B

I.D. No.	Quiz Number 1	2	3	4	3 Wk Test	6 Wk Test	Homework Average	Total
1	27	- 2	- 4	12	7	24	- 6	58
2	38	-15	- 2	7	21	25	10	84
6	- 5	-11	10	12	14	21	57	98
7	13	- 1	- 3	6	39	34	6	94
8	11	10	8	2	33	13	11	88
9	32	- 6	12	14	13	19	7	91
10	- 3	5	50	17	- 2	- 1	26	92
11	- 2	22	- 4	1	9	17	50	93
12	11	- 5	21	11	26	10	14	88
13	19	8	5	8	0	0	56	96
16	32	- 6	12	14	13	19	7	91
17	10	-10	12	4	45	5	21	87
18	34	-11	6	- 5	15	23	24	86
20	16	6	13	5	5	27	9	81
21	5	- 5	22	4	41	21	10	98
30	6	20	18	- 2	31	17	4	94
32	61	3	4	6	6	10	9	99
33	- 3	- 4	- 2	4	38	11	47	91
34	18	11	15	14	3	10	26	97
35	- 2	-13	1	2	19	2	48	57
36	39	- 8	23	7	6	19	- 6	80
40	23	-11	26	13	6	21	21	99
45	25	- 7	29	12	17	5	16	97
Mean	17.6	-1.3	11.8	7.3	17.6	15.3	20.3	88.6

This expectation is roughly borne out if we look at the column means for the quiz scores. The ordering of the four quizzes is exactly what would be predicted from the results of Form A, e.g. the first quiz in Form B accounts for an average 18 percent as compared with 22 percent (as the fourth quiz) in Form A, etc. One unexpected outcome is that for the second quiz where the average percent of variance is actually negative. This quiz consistently partials out variance in assigned marks when combined with the other six scores. This is probably related to the fact that the quiz has a high average correlation with the other six variables but has a small standard deviation. The author can give no satisfactory explanation of the difference between Form A and B for this quiz.

The variance attributable to the test, six-week exam, and homework average is also very similar to what would be expected even though many teachers gave differential weighting to these scores. The three-week test accounts for 18 percent (as opposed to 21 percent as the six-week test in A), the six-week test for 15 percent (as opposed to 11 as the homework average in A), and the homework accounts for 20 (compared with 16 as the three-week test in A). The ineffectiveness of differential weighting of raw scores is clearly illustrated by these data.

The average marks assigned on Form B correlate .25 with average marks assigned on Form A (N=23) and .01 with average final marks assigned for the second semester (N=18). There is a weak tendency for differences in standards to be consistent on the two exercises but no consistency between standards on the exercise and in the classroom. The latter result was also found with Form A.

Marking Questionnaire Data (Time 2)

The second administration of the Marking Questionnaire was handled in the same way as the administration of Form B of the Marking Exercise. Due to factors previously mentioned, the percent of completed questionnaires returned was disappointingly low.

A total of 28 (17 from School A and 11 from School B) completed questionnaires were returned and the response frequencies for each item were converted into proportions. These proportions are given item by item in Appendix B. The high attrition rate makes it impossible to draw direct comparisons between the Time 1 and Time 2 data. To facilitate such comparison the Time 1 data for only the 28 Time 2 respondents were used to compute new proportions. These data and a brief discussion of them are presented in Appendix B.

SUMMARY AND CONCLUSIONS

The data presented in this report were collected to answer two general questions. First, what is the nature and magnitude of individual differences among secondary school teachers with respect to:

(1) marks assigned on a variety of levels (daily work, quizzes, major tests, six-week and total semester reporting periods),

(2) marks assigned on a standardized marking exercise, and

(3) self-reported marking practices?

Second, to what extent are the individual differences mentioned above amenable to change as a result of participation in a concentrated in-service program on marking practices.

The first question is regarded as most crucial to an understanding of the marking process whereas the second is most relevant to an assessment of the effects of intervention upon this process. Two limitations of the study restrict the generalizability of the findings: (a) the sample consists of only 38 teachers (one class/teacher) from two different schools, and (b) no control group was employed. The absence of a control group means that statements concerning the impact of the in-service training experience are made on the assumption that no consistent shifts in marking practices occur between the first and second semester under "normal" circumstances.

The major findings are summarized below.

(1) The average marks on quizzes and daily work are higher than on six-week tests with marks

2

on both types of performance somewhat lower following the in-service training. The variability in average marks on quizzes and daily work is greater than on six-week tests with a slight increase in variability on both following the training.

(2) The average marks reported for the six-week and total semester periods are substantially higher than average semester examination marks both prior to and following the in-service program. The average reporting period and semester exam marks are slightly lower and more variable following the program.

(3) The correlation of average semester marks and average class IQ prior to the program is somewhat lower than following the program (.57 and .66, respectively). The correlation of average marks for the two different semesters is .82.

(4) The mean correlation between semester marks and IQ is the same before and after the program (.41 and .38, respectively) with great variation from class to class on both occasions. The correlations range from -.10 to .70 before and from -.15 to .90 after the training sessions.

(5) Approximately 95 percent of the variance in semester marks is accounted for by six-week and semester exam marks. Semester exam marks account for more variance than do six-week marks in most cases. Marks for the third six-weeks period generally account for more variance than do marks for the first and second periods. These results were found both before and after the in-service training. Great variation from teacher to teacher was also found (see Tables 15 and 22).

(6) Substantial inter-teacher agreement is found in the marks assigned on a standardized Marking Exercise. The intraclass correlation is .80 for Form A (prior to training) and .69 for Form B (following training). No statement of change is possible since the "pre" and "post"

samples were so different due to attrition on Form B. However, the range of marks assigned individual students is great on both Forms A and B (from F to C for the most extreme student on both forms).

(7) The average marks assigned on Form A are substantially higher than Form B (the latter having lower scores than the former) and the correlation of average marks assigned on the two forms of the standard exercise is .25. The correlations of these averages with average marks given in the classroom is substantially zero.

(8) The percent of variance in marks assigned on the standard exercises is generally predictable from the known characteristics (variances and intercorrelations) of the scores given in the exercises.

(9) There is great variation in self-reported marking practices both before and after the in-service training. There is no appreciable tendency towards greater concensus on most practices following the program.

The above statements provide a basis for several general conclusions about the marking process as it was studied and analyzed in this investigation.

First, the standards employed by teachers differ from one type of performance to the next with generally higher standards used on major tests and examinations than on daily work and quizzes. The standards which are employed appear to be subject to change although the change is not great in the present study.

Second, teachers show some sensitivity to general ability level within a class in assigning marks. This is manifested in a moderate positive relationship between average IQ and average mark. This sensitivity to between groups ability level appears to have been increased by the in-service program.

Third, despite the sensitivity to class ability level, a small part of the variance in marking standards appears to represent generalized marking habits (average marks assigned on different standard marking exercises show a low positive correlation) yet this does not generalize from an artificial situation to actual classroom marks.

Fourth, teachers assign marks largely in terms of absolute standards of performance and show little awareness of the relative difficulty level and discriminability of the performances upon which they base their marks. Since marks for the third six-weeks and semester exam typically have greater variability, these carry more weight in the determination of semester marks than do marks for the first and second six-weeks periods.

Fifth, most teachers apparently assign marks according to a global evaluative judgmental process which may or may not be related to objectively measured ability. The correlations among the marks assigned by an individual teacher for different periods and on the semester exam are uniformly high and yet the correlation of marks with IQ fluctuates radically from teacher to teacher. (This latter fluctuation can be only partially explained by differences in class ability levels.) The judgmental process by which marks are assigned is relatively immune to special training.

Finally, there are great individual differences in all facets of the marking process. These differences should not be ignored for they are often as important as the general findings outlined above.

References

Diederich, P. B., et al. "Factors in Judgments of Writing Ability." Res. Bull., 61-15, Educ. Test. Serv., Princeton, N. J., 1961.

Jansen, U. H. "Marking and Reporting Procedures in the Secondary Schools of Texas." Res. Bull., 32, Texas Study of Sec. Educ., 1960.

Johnson, F. W. The Administration and Supervision of a High School. Ginn and Co., 1925.

Morlan, G. "He Gave de Maupassant a Grade of B." Sch. and Soc., 34: 321-22, 1931.

Roelfs, R. M. Trends in Reporting to Parents on Pupil Progress in Junior High Schools. Unpublished Master's thesis, Univer. of Colorado, 1948, p. 90.

Roelfs, R. M. "Trends in Junior High School Progress Reporting." J. Ed. Res., 49: 241-9, 1955.

Starch, D. and Elliot, E. C. "The Reliability of Grading Work in English." Sch. Rev., 20: 442-57, 1912.

Starch, D. and Elliot, E. C. "The Reliability of Grading Work in Mathematics." Sch. Rev., 21: 254-59, 1913.

Terwilliger, James S. "Self-Reported Marking Practices and Policies in Public Secondary Schools." NASSP Bull., 50: 5-37, 1966.

Vroman, C., et al. "Preliminary Report of Survey of Grading Practices and Problems in North Central Association STS Project Schools." (Mimeo.) Ann Arbor, 1962.

Appendix A

Outline and Description of Materials Used in the In-Service Training Program

Outline of In-Service Training Program

Session I: Introduction

- A. Testing--First administration of Marking Questionnaire and distribution of Marking Exercise A
- B. General discussion of the nature and objectives of the in-service program

Session II. Importance of School Marks

- A. Necessity for evaluation
- B. Uses of marks
- C. Impact of marks upon the student
- D. Need to make marks relevant

Session III. Measurement and school marks

- A. Definition of measurement
- B. Special problems in psychological measurement
- C. Technical characteristics necessary in measurement
    - 1. standardization of procedure
    - 2. objectivity in scoring
    - 3. reliability
    - 4. validity
- D. Distinction between measurement and evaluation
- E. Need to base evaluation upon measurement

Session IV. Approaches to classroom measurement

- A. Types of format
    - 1. free-response items
    - 2. choice items
- B. Types of scoring
    - 1. objective
    - 2. subjective
- C. Purposes of classroom measurement
    - 1. diagnosis
    - 2. assessing mastery
    - 3. assessing general achievement
- D. General achievement tests and marking
- E. Non-test approaches
    - 1. classroom performance measures
    - 2. non-classroom performance

Session V. Describing achievement scores statistically

- A. Frequency distributions (construction and shapes)
- B. Percentiles and percentile ranks
- C. Central tendency (mean and median)
- D. Variability (Q and SD)
- E. Item analysis techniques
- F. Item characteristics and test scores

Session VI. Translating measures of achievement into marks

- A. Subject matter marks and non-achievement factors
- B. Objectivity and marks

Session VII. Different bases for marking

- A. Ability and achievement
- B. Self-improvement or growth
- C. The notion of absolute standards
- D. Achievement with reference to others

Session VIII. General achievement measures and variability

- A. Variability as an indicator of discrimination
- B. Variability as an index of relative weight when scores are added over tests

The in-service sessions lasted approximately one hour and were held in a classroom at each school immediately following the dismissal of school for the day. Detailed notes and supplementary materials for each session were typed, reproduced, and distributed to each participant. Each teacher was also given an inexpensive notebook cover in which a cumulative record of all the prepared materials was kept. Approximately 75 pages of such notes were given out during the eight sessions. In addition, all participants were given a complete Tests and Measurements Kit published by the Educational Testing Service. Selected readings on test construction and descriptive statistics were assigned in these references and discussed during the training sessions.

Appendix B

Marking Questionnaire
and Associated Data

The Marking Questionnaire was administered during the first in-service session and again during the final three weeks of the second semester. Table B.1 shows the percent responding to each option for all questions in Part II for the 41 teachers responding on the first administration and 2460 teachers in a previous national survey by Terwilliger (1966). The latter teachers are all teachers in the academic areas (Foreign Language, Language Arts, Mathematics, Science, and Social Studies) out of a total of 4000 teachers who responded in the survey. The survey data provides a convenient set of "norms" against which data from the present study may be compared. (Items where the percents do not sum to 100 were left blank by some teachers.)

There are no striking differences between the two sets of data if one considers the possibility of sampling error in the relatively small sample in the present study. There is great variability in the response to almost every item in both sets of data. The only issues on which there is considerable agreement concern the use of weighting formulas (over 90 percent don't use formulas) and the consideration given to final exams (most teachers endorsed "moderate consideration"). There is also general agreement among the 41 Nashville teachers on consideration given to work in previous marking periods (71 percent do not consider previous work at all) and relative consideration given to each marking period in assigning final marks (80 percent give equal consideration to all periods). These self-reports are interesting when compared with the results in Tables 15 and 17.

Table B 1

Percent Responding to Each Question on First Administration and for Comparison Data

	Time 1 (N=41)					Comparison (N=2460)				
Item	A	B	C	D	E	A	B	C	D	E
9	46	29	15	5	5	28	38	27	7	0
10	66	29				57	43			
11	2	98				8	92			
12	20	44	31	5		29	47	21	4	
13	8	46	24	15		21	44	25	9	
14	7	37	39	15		5	31	44	19	
15	2	22	52	22		6	30	48	15	
16	2	27	49	17		7	14	50	29	
17	0	0	30	68		6	1	22	76	
18	5	22	37	34		5	35	42	18	
19	71	17	7	0		53	27	12	8	
20	0	2	93	5		4	9	63	23	
21	17	17	44	22		10	37	43	10	
22	15	49	31	5		(22-24 not included				
23	49	32	17	0		(in survey question-				
24	2	10	80	6		(naire				

Table B 2 gives comparable item response data for the 28 teachers who answered the questionnaire on both administrations. The left side of the table gives time 1 data and the right side, time 2. In general, the results are highly similar on the two administrations. The only items suggesting change are those dealing with homework (less consideration, time 2), term papers, etc. (more consideration, time 2), and standardized test results (less consideration, time 2). None of these changes can be attributed to the program because these issues were not dealt with directly in the training sessions.

Table B 2

Percent Responding to Each Question for Only Subjects Who Completed the Questionnaire Time 1 and 2 (N=28)

	Time 1					Time 2				
Item	A	B	C	D	E	A	B	C	D	E
9	46	27	18	9	0	50	25	21	4	0
10	64	27				71	25			
11	9	91				11	89			
12	27	46	27	0		28	61	11	0	
13	9	46	27	9		32	46	22	0	
14	9	46	27	9		11	46	39	4	
15	0	27	27	37		4	28	57	11	
16	0	18	65	18		0	21	61	18	
17	0	0	9	83		0	4	18	78	
18	0	37	27	27		0	43	46	11	
19	73	9	9	0		78	14	8	0	
20	0	0	100	0		0	0	75	25	
21	27	27	37	9		7	32	50	11	
22	27	55	18	0		21	54	25	0	
23	55	45	0	0		82	18	0	0	
24	0	9	91	0		0	4	96	0	

MARKING QUESTIONNAIRE

Part I

Identifying Information

1. Name____ 2. Sex______

3. Highest Degree __________ 4. Total Years Teaching Experience_____

5. Indicate the general subject area of the first class which you teach on Monday mornings.

   a. Foreign Language

   b. Language Arts

   c. Mathematics

   d. Science

   e. Social Studies

   f. Other__________________

6. Give the exact course title in the space below.

7. Grade level(s) of the course_____________________

8. Number of years you have taught this course______

Part II

Marking Practices

The following questions are designed to obtain information about practices you employ in assigning marks to students. We specifically want to know about practices used in the course which you named on the previous page. Please keep that course in mind as you answer the following questions. Respond by circling the letter corresponding to the answer which is most descriptive for you. (Circle only one answer for each item.)

9. Which of the following best represents the primary basis which you use in determining the marks of your students?

   a. absolute achievement
   b. achievement with respect to ability
   c. achievement with respect to the class
   d. self-improvement or growth
   e. other (please specify)

10. Do you use some kind of formula to weight the factors (homework, tests, etc.) which you consider in assigning marks?

    a. yes
    b. no
    If yes, please write the formula.

11. Do you assign marks to approximate some predetermined distribution, e.g., 5% A's, 15% B's, etc.?

    a. yes
    b. no
    If so, please describe the distribution you use.

Questions 12-19 concern consideration you give to various factors in determining marks for each regular six weeks marking period.

12. What consideration do you give to classroom behavior (e.g., causing distractions or conversely being very cooperative) in determining course marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

13. What consideration do you give to unexcused absence and tardiness in determining marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

14. What consideration do you give to "effort" in assigning marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

15. What consideration do you give to homework in determining marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

16. What consideration do you give quiz (less than 15 minutes) scores in determining marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

17. What consideration do you give to test (up to one class period) scores in assigning marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

18. What consideration do you give to classroom performance (i.e., recitation, voluntary participation, etc.) in determining marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-99%)

19. What consideration do you give to work done in previous marking periods when assigning marks for a period just ended (e.g., to what extent do marks given for the second period represent work done during the first period)?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-100%)

Questions 20-24 refer to consideration you give to factors in determining final semester marks.

20. What consideration do you give to final examinations (end of semester) in determining final marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-100%)

21. What consideration do you give to other assigned work (term papers, individual projects, etc.) in determining final marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-100%)

22. What consideration do you give to "make-up" or "extra-credit" work voluntarily done by students?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-100%)

23. What consideration do you give to standardized test results (ability or achievement measures) in determining final marks?

    a. no consideration at all (0%)
    b. minor consideration (1-10%)
    c. moderate consideration (11-25%)
    d. major consideration (26-100%)

24. Indicate the relative consideration which is given to work done in each marking period when you assign final marks.

    a. more consideration is given to work done in earlier periods (e.g., first six weeks)
    b. more consideration is given to work done in later periods (e.g., last six weeks)
    c. equal consideration is given to work done in all marking periods.
    d. other (specify)

Part III

Attitudes Concerning Marking

The statements in this section represent opinions and attitudes about marks and marking which are often expressed by teachers and students. You are to indicate the degree to which you agree or disagree with each statement by circling one of five alternatives:

SA = strongly agree

A = agree

U = uncertain

D = disagree

SD = strongly disagree

There are no right or wrong answers. We simply want your reaction to each statement. Please try to give a response to each statement. Do not omit an item unless you feel that it is impossible to give a meaningful answer.

SA A U D SD 25. Students who are not concerned about marks are unrealistic.

SA A U D SD 26. The most important factors in determining marks can only be judged subjectively.

SA A U D SD 27. Objective type examinations accurately assess the more important types of learning.

SA A U D SD 28. Marks should be based upon fixed standards which a teacher maintains irrespective of how students perform.

SA A U D SD 29. School marks represent a system of reward and punishment which does more harm than good.

SA A U D SD 30. "Social promotion" is more desirable than failing students with limited ability.

SA A U D SD 31. A teacher should not consider other demands on student time (involvement in athletics, school organizations, after-school jobs, etc.) when assigning marks.

SA A U D SD 32. Students should not be allowed to do special make-up extra-credit work in order to raise their marks.

SA A U D SD 33. Students and their parents attach too much importance to marks.

SA A U D SD 34. "Effort" should not be considered in assigning marks to students.

SA A U D SD 35. The more important achievements of students can be measured quantitatively.

SA A U D SD 36. More important types of learning can be evaluated only by subjectively scored (essay) tests.

SA A U D SD 37. Marks should be based upon the relative levels of achievement among students in a class.

SA A U D SD 38. Course marks should not be lowered for disciplinary reasons.

SA A U D SD 39. A teacher should consider student home environment (parents income and education, whether parents are divorced, etc.) when assigning marks.

SA A U D SD 40. Assigning marks is one of the most distasteful aspects of teaching.

SA A U D SD 41. What a student learns is more important than the mark he makes.

SA A U D SD 42. The assignment of marks to students is, at best, a necessary evil.

SA A U D SD 43. Differences in the natural ability of students should be considered in assigning school marks.

SA A U D SD 44. A good teacher can tell how much a student has learned without giving written tests.

SA A U D SD 45. Classroom deportment (causing distractions, etc.) should be reflected in subject matter marks.

SA A U D SD 46. Marks should be assigned to approximate some predetermined distribution, e.g., 5% *A*'s, 15% *B*'s, etc.

SA A U D SD 47. Competition among students for good marks is a healthy thing.

SA A U D SD 48. The more important outcomes of an education are reflected by school marks.

SA A U D SD 49. Our schools would not be better off if marks were abolished.

SA A U D SD 50. Marks should be based primarily upon the teacher's evaluation of student progress or growth.

SA A U D SD 51. The contributions a student makes in the classroom through discussion and *voluntary* participation should not be considered in assigning marks.

SA A U D SD 52. When a student fails a subject it is as much a failure of the teacher as the student.

SA A U D SD 53. Some students may fail in spite of the quality of the teaching.

SA A U D SD 54. Marks should be based solely upon teacher-made tests and other written work.

Appendix C

Marking Exercises
A and B

The two forms of the Marking Exercise employed in the present study were designed as vehicles for the investigation of marking practices under well-controlled and standardized conditions. Every attempt was made to make the exercises both realistic and interesting to the teachers.

Form A was designed several years ago and has been used in other research. The biographical data developed for each student were designed to give a life-like, integrated and consistent description. The classroom quiz, test, and homework scores were also structured so as to be both realistic and consistent. The means, standard deviations, and correlations among the seven classroom performance scores are given below.

Table C 1

	Quiz Number				3 Wk Test	6 Wk Test	Homework Average
	1	2	3	4			
Q 1	---						
u 2	.53	---					
i 3	.60	.45	---				
z 4	.54	.61	.62	---			
3 Wk Test	.30	.36	.75	.67	---		
6 Wk Test	.59	.49	.61	.59	.41	---	
Homework Average	.61	.41	.66	.43	.55	.38	---
Mean	78.4	69.8	89.2	70.0	79.9	86.4	92.8
SD	13.4	11.3	3.8	19.0	9.6	6.6	4.6

Form B was derived from Form A in the following way:

1. names and the wording of the biographical descriptions were changed but the sex, IQ, and general personality traits of each student were retained
2. slight modifications were made in the absence and tardiness figures
3. quiz scores were altered and/or shifted as below
    a. 1st week quiz scores were shifted to the 4th week quiz
    b. five was subtracted from each 2nd week quiz score and these scores were shifted to the 5th week quiz
    c. four was subtracted from each 4th week quiz score and these scores were shifted to the 2nd week quiz
    d. six was subtracted from each 5th week quiz score and these scores were shifted to the 1st week quiz
4. test, exam and homework average scores were altered and/or changed as below
    a. three-week test scores were shifted to the homework average
    b. ten was subtracted from each six-week exam score and these scores were shifted to the three-week test
    c. eight was subtracted from each homework average and these scores were shifted to the six-week exam

The alterations involving the quiz, test, six-week exam, and homework data systematically reduce the means by the constants subtracted but, of course, change neither the variances nor the relationships among the variables. Shigting the variables about simply yields a different permutation of variances for specific quizzes and tests, e.g. the 1st quiz is now the variable with the greatest variance. Therefore, the means, standard deviations, and correlations given for Form A also describe Form B if appropriate changes in means and variable names are made.

Marking Exercise

The attached pages contain brief biographical descriptions of twenty hypothetical students. These descriptions include each student's name, an I.Q. score (in parentheses following name) and selected information concerning classroom behavior, general academic performance and extra-curricular activities. Miscellaneous observations are also provided for most students.

These twenty students have just completed a six-weeks period in a class together. The final page contains attendance data and quiz, examination and home work marks for each student. Weekly quizzes were given at the end of the first, second, fourth and fifth weeks. A mid-period test was given at the end of the third week and a comprehensive examination was given at the end of the six-weeks period. The home work mark is an average for all work during the six weeks period. The marks which have been recorded are based upon a 0-100 scale.

You are to assign six weeks marks to each student using both a 0-100 scale and a marking system used by your own school. (For example, if you assign a student a mark of 94 and 93-100 is the range designated as "A" at your school, you would also give the student an "A.")

Mark these students just as you would mark students in your own classes insofar as this is possible with the information which you have been given.

1. William Allen (120)

Classroom Behavior: Steady performance in subjects which interest him. Almost no effort expended in less interesting subjects.

General Academic Performance: Grades range from A's thru C's and an occasional D.

Extracurricular: Reporter, School Newspaper, Track Team, Glee Club.

Miscellaneous: An aggressive personality--"a fighter." A person of small build who makes up for it in a "big" attitude.

2. Winifred Boynton (110)

Classroom Behavior: Well prepared to participate in class--adds considerably to class discussion and is attentive listener as well.

General Academic Performance: A very serious student with a straight "A" average.

Extracurricular: No outside activities--all effort is expended in carrying out assignments and doing extra work in related subjects.

Miscellaneous: Has few if any friends. An introverted personality. Peers characterize her as an "egghead."

3. Cathryn Brinckerhoff (126)

Classroom Behavior: Strong performer in all classes. Well motivated. Active class participation.

General Academic Performance: Ranks in top 10th of class of 200 students.

Extracurricular: Active participator in extracurricula activities.

Editor--Yearbook
National Society
Class Plays
NFL

Miscellaneous: Respected by peers but not really well liked by many because of a superiority complex. Teachers consider student an asset to their classes. Student is intelligent, capable; grades come easily.

4. Edward Dent (99)

Classroom Behavior: Seems to be interested in the class content and participates but something is keeping this student from achieving to his full capacity.

General Academic Performance: Grades in low B's and C's.

Extracurricular: Track Team
Swimming Team

Miscellaneous: Good personality, outgoing and liked by everyone.

5. Patricia Fischer (107)

Classroom Behavior: Characterized by a wide variety of interests as well as high motivation toward academic work.

General Academic Performance: A low "A" student--some "B's".

Extracurricular: Secretary-Student Council
Queen-Junior-Senior Prom
YWCA Committee
Glee Club

Miscellaneous: A very attractive girl who's "brainy" to boot. Warm outgoing personality equally liked by students and teachers. Has ability to conceal intellectual talents when appropriate.

6. Doris Gulick (111)

Classroom Behavior: Moderate participation in class; sometimes appears disinterested.

General Academic Performance: Grades are divided equally among B's and C's.

Extracurricular: Cheer Leader
Homecoming Queen
Dramatics

Miscellaneous: Student is more interested in having a good time than in serious study. Student is very attractive, has entered several beauty contests. Very popular with the student body. Liked by teachers.

7. Elizabeth Henderson (116)

Classroom Behavior: Is interested enough in work to want to build a good achievement record.

General Academic Performance: Grades mostly A's and B's. Grades come easily in some areas but student puts forth considerable work in other subjects.

Extracurricular: Glee Club
Yearbook Staff

Miscellaneous: Has been in and out of school intermittenly during the past year due to serious illness. Has attempted to keep up with work at home and has done rather well in this attempt. Has had to curb extracurriculars.

8. Nancy Jennings (94)

Classroom Behavior: What is learned ais gained thru classroom participation. Student is well motivated but has little time to spend in outside preparation due to home situation.

General Academic Performance: Grades mostly B's with a scattering of C's. Very little outside study.

Extracurricular: None

Miscellaneous: Mother an invalid--father works nights. Student has responsibility for care of five brothers and sisters. School activities also limited due to this factor. Student is mature for age. Has no close circle of friends.

9. Richard Kim (92)

Classroom Behavior: Must work hard for everything he gets but is sufficiently interested in content to put forth the needed effort. Has very good organizing ability which helps him to deal with course content.

General Academic Performance: Strong B student.

Extracurricular: Editor, School Newspaper, Track Team

Miscellaneous: Well liked by peers--a lanky, outgoing person with a serious side to his personality.

10. Frederick Lynch (132)

Classroom Behavior: An intelligent but reserved student. Very little active participation in class discussion but attentive to what is taking place. Occasionally will make a good point in class discussion.

General Academic Performance: Grades mostly A's with some B's.

Extracurricular: Baseball Team
Track Team
National Honor Society

Miscellaneous: Quite reserved except within a close circle of friends.

11. Irene Mason (93)

Classroom Behavior: Student is more anxious to be heard than to understand what is going on. Has an opinion on everything. Does little synthesizing and there is some question as to how much material has really sunk in.

General Academic Performance: Primarily a B student with some C's.

Extracurricular: Debate Team
Speech Team
Student Council
Assembly Chairman

Miscellaneous: A rather unattractive girl who hasn't learned "silence is golden."

12. Kenneth Michael (104)

Classroom Behavior: Seems to be completely unaware of what is taking place in class. Never adds to class discussion. Disinterested.

General Academic Performance: Grades mostly C's and D's with an occasional F. Is currently repeating this course due to previous failure.

Extracurricular: None

Miscellaneous: Has been in trouble with school authorities and is presently on probation. Associates with an older working group of boys; has been involved in some incidents with authorities. Comes from broken home.

13. James Newtton (91)

Classroom Behavior: Distracts from classroom order by carrying on private conversations with classmates. Is not motivated except in rare instances. Then exhibits a spirit of activity only to fall back into the old pattern.

General Academic Performance: Grade range, B's and C's.

Extracurricular: Has ability beyond achievement. Has some interest in basketball but does not participate actively in any organization to which he belongs.

Miscellaneous: Has been in trouble with school authorities at times but is not on probation at the moment.

14. Dennis O'Connor (101)

Classroom Behavior: A very shy student who never participates in class discussion. Is alert and attentive but never voices his own thoughts or opinions.

General Academic Performance: Grades are about evenly distributed among B's and C's.

Extracurricular: Interested in writing and is a frequent contributor to the school literary magazine. Participates in no team sports or group activities.

Miscellaneous: Student is very conscious of a slight speech defect.

15. David Peterson (128)

Classroom Behavior: Performance even in all classes-well motivated.

General Academic Performance: Straight A average.

Extracurricular: School Debating Team
National Honor Society
National Forensic League
Contributor, School Literary Magazine
Member, Student Council
Photographer, Yearbook

Miscellaneous: A model student, well liked by teachers and small circle of friends.

16. Edgar Phillips (84)

Classroom Behavior: Tries hard but cannot gain better grasp of class material. Is eager to add to class discussion but cannot grasp the central core of ideas and only repeats what has already been said.

General Academic Performance: Straight C Average.

Extracurricular: Football Team

Miscellaneous: Nearly the largest member of the class but well liked by peers. Easy going personality.

17. Paula Ranalle (96)

Classroom Behavior: A moderately motivated student. One who is more interested in social relations with peers than in a serious study.

General Academic Performance: Grades in B's and C's.

Extracurricular: Very good in dramatics and quite interested in this area.

Miscellaneous: Has a good personality and is a non-conformist in a way--the school "character". Well liked in the school.

18. Judy Strang (97)

Classroom Behavior: Student's attention is never concentrated on the moment.

General Academic Performance: Primarily a B student with some C's and some A's. Has quite a bit of native ability and interests are widely varied.

Extracurricular: Dramatics
President, Chemistry Club
Musical Combo
Student Council
Yearbook Staff

Miscellaneous: Impresses others as being scatterbrained; so many things going at once. Student is quite small and always appears busy.

19. Samuel Thorp (104)

Classroom Behavior: Student makes little impression on anyone. Difficult to say what makes him "tick." Is a completely neutral person.

General Academic Performance: Grades range from A's thru C's, mostly B's.

Extracurricular: Participates in no outside activities due to an after school job.

Miscellaneous: None

20. Jean Tindall (130)

Classroom Behavior: Often makes a cogent point during class discussion.

General Academic Performance: Grades are mostly A's with a scattering of B's. Is eager to learn and has ability to synthesize material learned in different classes.

Extracurricular: Dramatics
Future Teachers of America
Cheer Leader
Yearbook Staff
Class Activities

Miscellaneous: Student has a genuine like for people and tries hard to make friends. Peers sometimes shy away because they feel keenly this student's native ability to be much above theirs.

	Times Absent	Times Tardy	Weekly Quiz Scores				Three Weeks Test	Six Weeks Exam	Home Work Average	Six Weeks Grade	
			1st wk.	2nd wk.	4th wk.	5th wk.				No.	Letter
1. William Allen	3	1	81	67	88	64	86	90	93		
2. Winifred Boynton	1	0	80	86	96	100	98	95	95		
3. Cathryn Brinckerhoff	0	1	100	67	93	99	90	93	100		
4. Edward Dent	0	0	94	65	89	51	76	88	94		
5. Patricia Fischer	0	1	89	71	94	73	80	87	100		
6. Doris Gulick	5	0	72	70	96	78	95	86	92		
7. Elizabeth Henderson	8	0	80	71	90	53	80	86	90		
8. Nancy Jennings	7	1	90	90	88	62	60	91	92		
9. Richard Kim	0	0	67	62	88	58	72	88	93		
10. Frederick Lynch	0	0	86	100	90	100	90	91	95		
11. Irene Mason	1	1	79	61	86	67	72	74	86		
12. Kenneth Michael	10	7	50	58	84	50	78	74	91		
13. James Newtton	1	1	73	65	85	67	68	79	86		
14. Dennis O'Conner	0	0	55	66	83	67	76	82	88		
15. David Peterson	2	0	100	79	91	97	85	98	96		
16. Edgar Phillips	2	0	65	54	89	57	75	93	88		
17. Paula Ranallo	2	0	67	57	85	68	70	85	85		
18. Judy Strang	1	2	72	62	87	39	76	82	91		
19. Samuel Thorp	3	0	74	67	88	50	78	76	100		
20. Jean Tindall	1	0	[illegible]	77	95	100	95	90	100		

Appendix D

Multiple Regression Analyses for Six-Week Marking Periods

A special multiple regression analysis was performed on the data of five teachers selected from each school. The purpose of this analysis was to determine the relative weight carried by different types of data in the determination of six-week marks. The six-week marks for each marking period during each semester were predicted from all other data collected from the record book for the six-week period. The predictors in most instances were three quiz or daily work grades, a six-week daily average grade, and a six-week test grade. The results are shown in Table D 1.

Each row of the table represents a six-week period with first semester data on the left and second semester on the right. The percents of variance accounted for by the three quiz or daily marks are given in the columns headed "Q 1, Q 2, and Q 3"; the percent attributable to daily average under the column headed "DA"; and the percent attributable to the six-week test in the column labeled "6 WT". A dash indicates that a teacher did not record marks for a specific predictor score.

Where all five predictors are available, the total percent of variance accounted for is consistently high for most teachers. This implies that the predictor variables used in most analyses represent the important factors upon which teachers based their six-week marks. Of course, this is to be expected because the daily average and six-week test are included in the predictor variables. The exceptions are Teachers 14 (second semester), 31 (second semester), and 46 (both semesters). There is no apparent explanation for these low values so it seems appropriate to conclude that the six-week marks in these cases are based to a substantial degree upon subjective judgments unrelated to data recorded by the teacher.

Table D 1

Percent of Variance in Six-Weeks Marks Attributable to Quiz, Daily Average, and Six-Weeks Test Marks

I.D. No.	6 Wks	First Semester						Second Semester					
		Q1	Q2	Q3	DA	6 WT	To-tal	Q1	Q2	Q3	DA	6 WT	To-tal
4	1	-14	07	13	32	59	97	6	8	2	48	33	97
	2	-2	1	9	56	32	96	-1	10	13	44	19	85
	3	1	-4	-3	61	42	97	0	0	0	100	0	100
8	1	5	10	-2	--	78	91	-1	1	0	35	64	99
	2	-4	3	4	31	62	96	3	0	0	22	74	99
	3	0	40	0	--	--	40	0	0	0	0	100	100
11	1	19	7	16	23	31	96	-3	-3	7	16	56	73
	2	17	7	28	2	40	94	6	11	-1	62	21	99
	3	6	11	-5	34	43	89	-13	-1	1	110	--	97
12	1	-5	-5	2	98	9	99	2	8	6	56	22	94
	2	-1	5	0	54	37	95	5	-5	-3	58	44	99
	3	-4	3	-3	69	28	93	3	1	7	87	--	98
14	1	-2	1	-2	62	40	99	-11	24	5	55	-11	62
	2	2	-3	-9	63	45	98	0	-1	0	52	38	89
	3	16	7	29	-6	37	83	9	8	2	28	33	80
31	1	15	14	8	--	56	93	1	-17	26	37	25	72
	2	7	16	3	--	65	91	27	11	-3	75	-34	76
	3	4	45	36	--	--	85	19	34	6	--	22	81
32	1	3	2	1	39	53	98	-1	-7	-2	59	43	92
	2	-2	3	2	48	47	98	-3	-1	-2	57	45	96
	3	56	18	2	--	--	76	-2	8	4	55	--	65
39	1	0	0	0	100	0	100	0	0	0	0	100	100
	2	0	0	0	100	0	100	0	0	0	--	100	100
	3	0	0	0	100	0	100	0	6	--	--	--	6
45	1	-10	-2	-30	51	88	97	12	6	--	66	13	97
	2	-17	5	-3	86	23	94	7	-3	8	79	8	99
	3	3	-5	10	58	31	97	-1	-8	--	57	47	95
46	1	0	1	0	34	37	72	11	24	11	--	35	81
	2	2	6	-1	54	8	69	5	4	10	40	30	89
	3	1	7	45	16	3	72	40	8	6	--	--	54

One highly unusual outcome is that for the third six-weeks of the second semester for Teacher 11. The daily average accounts for more than 100 percent because it correlates .98 with the criterion and its beta weight exceeds 1.00. Two of the quizzes have negative beta weights so the total variance accounted for is less than 100 percent.

There are several instances where one predictor accounts for all of the variance in six-week marks, viz. third six-weeks of second semester for Teachers 4 and 8 and all periods except third six-weeks of second semester for Teacher 39. The predictor which correlates perfectly with marks is either daily average or the six-week test. Therefore, the teacher simply used the daily average (or the six-week test) grade as the grade for the six-week period despite the fact that quiz and six-week test grades (or quiz and daily average grades) were available. At the other extreme, we find that the marks for the third six-weeks of semester one for Teacher 8 and the third six-weeks of semester two for Teacher 39 cannot be accounted for to any appreciable extent. Neither daily average nor six-week test marks had been recorded and only one of the available quiz marks accounts for any variance. It is impossible to say what the teacher based six-week marks upon under these circumstances unless it was subjective judgment.

The six-week test and daily average, when recorded, account for more variance than the quizzes for most teachers. However, there appears to be no consistency in the relative importance of the quizzes, daily average, and test for a given teacher over different six-week periods. This reflects differences in the magnitude of the predictor and criterion variances and intercorrelations from one period to the next. Also, since most teachers reported that they do not use a formula in assigning marks (see Appendix B) a fair amount of subjective judgment undoubtedly affects the weights carried by the various predictors. The systematic use of simple weighting formulas would tend to stabilize the relative importance of the predictors and would also give both the teachers and students a better understanding of marks.

Appendix E

Mean IQ and Correlations of IQ and Reported Marks for Individual Classes

TEST-GRADE CORRELATIONS
First Semester

Teacher 1 N=24 Biology Mean IQ=107

Otis	---					
First 6 Wks.	529	---				
Second 6 Wks.	471	895	---			
Third 6 Wks.	473	874	859	---		
Sem. Ex.	433	831	841	817	---	
Sem. Mk.	457	924	933	939	929	---

Teacher 2 N=25 American Government Mean IQ=118

ACT	---						
Otis	910	---					
First 6 Wks.	669	544	---				
Second 6 Wks.	733	607	640	---			
Third 6 Wks.	723	496	658	658	---		
Sem. Ex.	578	449	700	722	744	---	
Sem. Mk.	731	543	852	805	876	878	---

Teacher 3 Data not available.

Teacher 4 N=19 Spanish II Mean IQ=115

Otis	---					
First 6 Wks.	238	---				
Second 6 Wks.	300	882	---			
Third 6 Wks.	036	737	741	---		
Sem. Ex.	207	815	869	844	---	
Sem. Mk.	256	915	924	894	928	---

Teacher 5 Less than 15 students with complete data.
Teacher 6 Less than 15 students with complete data.

Teacher 7 N=22 English III Mean IQ=124

Otis	---					
First 6 Wks.	366	---				
Second 6 Wks.	109	467	---			
Third 6 Wks.	198	615	597	---		
Sem. Ex.	446	633	481	855	---	
Sem. Mk.	389	736	683	925	887	---

Teacher 8 N=24 English IV Mean IQ=118

ACT	---						
Otis	637	---					
First 6 Wks.	405	215	---				
Second 6 Wks.	186	036	751	---			
Third 6 Wks.	252	177	790	578	---		
Sem. Ex.	412	319	707	726	578	---	
Sem. Mk.	378	258	917	820	788	888	---

Teacher 9 N=18 World History Mean IQ=108

Otis	---					
First 6 Wks.	620	---				
Second 6 Wks.	647	924	---			
Third 6 Wks.	631	893	926	---		
Sem. Ex.	808	803	815	901	---	
Sem. Mk.	671	941	967	973	909	---

Teacher 10 N=16 Spanish I Mean IQ=116

Otis	---					
First 6 Wks.	304	---				
Second 6 Wks.	216	774	---			
Third 6 Wks.	189	851	814	---		
Sem. Ex.	124	585	781	699	---	
Sem. Mk.	234	812	836	874	785	---

Teacher 11 N=21 Biology Mean IQ=116

Otis	---					
First 6 Wks.	420	---				
Second 6 Wks.	372	867	---			
Third 6 Wks.	542	919	864	---		
Sem. Ex.	491	839	909	853	---	
Sem. Mk.	494	878	927	895	945	---

Teacher 12 N=26 Latin II Mean IQ=121

Otis	---					
First 6 Wks.	523	---				
Second 6 Wks.	481	797	---			
Third 6 Wks.	638	842	914	---		
Sem. Ex.	674	654	733	826	---	
Sem. Mk.	603	840	922	995	837	---

Teacher 13 N=23 Algebra I Mean IQ=112

Otis	---					
First 6 Wks.	028	---				
Second 6 Wks.	018	649	---			
Third 6 Wks.	-281	480	658	---		
Sem. Ex.	039	210	515	459	---	
Sem. Mk.	044	589	842	821	706	---

Teacher 14 N=24 English II Mean IQ=113

Otis	---					
First 6 Wks.	511	---				
Second 6 Wks.	622	786	---			
Third 6 Wks.	497	574	800	---		
Sem. Ex.	500	765	844	554	---	
Sem. Mk.	630	858	952	832	856	---

Teacher 15 N=23 Algebra I Mean IQ=124

Otis	---					
First 6 Wks.	-040	---				
Second 6 Wks.	-219	749	---			
Third 6 Wks.	-170	757	830	---		
Sem. Ex.	067	330	486	649	---	
Sem. Mk.	-117	761	875	951	762	---

Teacher 16 N=25 Sociology Mean IQ=116

ACT	---						
Otis	808	---					
First 6 Wks.	495	448	---				
Second 6 Wks.	456	324	650	---			
Third 6 Wks.	324	077	538	753	---		
Sem. Ex.	521	515	820	645	438	---	
Sem. Mk.	552	437	852	900	784	838	---

Teacher 17 N=26 American History Mean IQ=121

Otis	---					
First 6 Wks.	-164	---				
Second 6 Wks.	082	533	---			
Third 6 Wks.	153	038	312	---		
Sem. Ex.	-031	-162	260	571	---	
Sem. Mk.	-034	091	494	767	879	---

Teacher 18 Less than 15 students with complete data.

Teacher 19 N=28 Algebra II Mean IQ=112

Otis	---					
First 6 Wks.	171	---				
Second 6 Wks.	253	779	---			
Third 6 Wks.	158	701	652	---		
Sem. Ex.	135	633	599	566	---	
Sem. Mk.	230	881	866	831	827	---

Teacher 20 N=23 Biology Mean IQ=108

Otis	---					
First 6 Wks.	631	---				
Second 6 Wks.	532	700	---			
Third 6 Wks.	735	825	789	---		
Sem. Ex.	556	628	743	646	---	
Sem. Mk.	654	860	927	894	853	---

Teacher 21 Data not available.

Teacher 30 N=19 Advanced Math Mean IQ=119

Otis	---					
First 6 Wks.	085	---				
Second 6 Wks.	348	781	---			
Third 6 Wks.	261	782	766	---		
Sem. Ex.	217	710	742	672	---	
Sem. Mk.	239	849	857	853	923	---

Teacher 31 N=24 Unified Geometry Mean IQ=107

Otis	---					
First 6 Wks.	554	---				
Second 6 Wks.	597	746	---			
Third 6 Wks.	660	755	841	---		
Sem. Ex.	550	552	636	634	---	
Sem. Mk.	626	813	908	910	778	---

Teacher 32 N=23 Biology Mean IQ=109

Otis	---					
First 6 Wks.	528	---				
Second 6 Wks.	563	877	---			
Third 6 Wks.	597	776	881	---		
Sem. Ex.	676	839	829	816	---	
Sem. Mk.	661	921	946	885	920	---

Teacher 33 Less than 15 students with complete data.
Teacher 34 Less than 15 students with complete data.

Teacher 35 N=24 Biology Mean IQ=104

Otis	---					
First 6 Wks.	611	---				
Second 6 Wks.	496	565	---			
Third 6 Wks.	528	723	755	---		
Sem. Ex.	581	408	638	808	---	
Sem. Mk.	627	667	833	951	897	---

Teacher 36 N=30 World Geography Mean IQ=106

Otis	---					
First 6 Wks.	588	---				
Second 6 Wks.	502	804	---			
Third 6 Wks.	407	653	709	---		
Sem. Ex.	731	745	576	555	---	
Sem. Mk.	665	823	782	817	861	---

Teacher 37 N=21 Sociology Mean IQ=105

Otis	---					
First 6 Wks.	428	---				
Second 6 Wks.	233	861	---			
Third 6 Wks.	010	778	781	---		
Sem. Ex.	367	866	810	844	---	
Sem. Mk.	261	919	885	918	953	---

Teacher 38 N=17 World History Mean IQ=105

Otis	---					
First 6 Wks.	201	---				
Second 6 Wks.	370	519	---			
Third 6 Wks.	219	665	531	---		
Sem. Ex.	419	766	652	698	---	
Sem. Mk.	352	889	711	865	911	---

Teacher 39 N=21 American History Mean IQ=113

Otis	---					
First 6 Wks.	222	---				
Second 6 Wks.	141	458	---			
Third 6 Wks.	137	647	793	---		
Sem. Ex.	271	482	716	699	---	
Sem. Mk.	289	597	853	925	824	---

Teacher 40 N=20 Algebra I Mean IQ=101

Otis	---					
First 6 Wks.	579	---				
Second 6 Wks.	699	742	---			
Third 6 Wks.	636	768	882	---		
Sem. Ex.	610	729	787	807	---	
Sem. Mk.	682	828	933	948	912	---

Teacher 41	Data not available.
Teacher 42	Data not available.
Teacher 43	Less than 15 students with complete data.

Teacher 45 N=21 Spanish I Mean IQ=117

Otis	---					
First 6 Wks.	262	---				
Second 6 Wks.	370	770	---			
Third 6 Wks.	269	763	900	---		
Sem. Ex.	325	735	747	844	---	
Sem. Mk.	362	840	892	935	934	---

Teacher 46 N=25 English II Mean IQ=107

Otis	---					
First 6 Wks.	385	---				
Second 6 Wks.	527	773	---			
Third 6 Wks.	397	796	623	---		
Sem. Ex.	287	668	422	459	---	
Sem. Mk.	322	943	767	771	702	---

Teacher 47 Less than 15 students with complete data.

Teacher 48 N=22 Latin I Mean IQ=112

Otis	---					
First 6 Wks.	424	---				
Second 6 Wks.	383	915	---			
Third 6 Wks.	467	742	722	---		
Sem. Ex.	411	667	705	768	---	
Sem. Mk.	493	878	898	895	871	---

TEST-GRADE CORRELATIONS
Second Semester

Teacher 1 N=24 Biology Mean IQ=107

Otis	---					
First 6 Wks.	503	---				
Second 6 Wks.	429	771	---			
Third 6 Wks.	438	829	853	---		
Sem. Ex.	382	661	587	668	---	
Sem. Mk.	422	880	791	880	863	---

Teacher 2 Data not available.

Teacher 3 N=21 English I Mean IQ=114

Otis	---					
First 6 Wks.	151	---				
Second 6 Wks.	221	493	---			
Third 6 Wks.	-051	502	416	---		
Sem. Ex.	006	559	444	722	---	
Sem. Mk.	066	684	662	829	916	---

Teacher 4 N=19 Spanish II Mean IQ=115

Otis	---					
First 6 Wks.	197	---				
Second 6 Wks.	171	659	---			
Third 6 Wks.	-141	595	620	---		
Sem. Ex.	072	789	621	706	---	
Sem. Mk.	079	783	774	867	908	---

Teacher 5 Less than 15 students with complete data.
Teacher 6 Less than 15 students with complete data.

Teacher 7 N=23 English III Mean IQ=124

Otis	---					
First 6 Wks.	112	---				
Second 6 Wks.	089	734	---			
Third 6 Wks.	172	306	228	---		
Sem. Ex.	134	345	352	864	---	
Sem. Mk.	155	574	501	876	888	---

Teacher 8 N=24 English IV Mean IQ=118

ACT	---						
Otis	637	---					
First 6 Wks.	332	097	---				
Second 6 Wks.	445	197	805	---			
Third 6 Wks.	371	156	709	913	---		
Sem. Ex.	543	375	724	786	746	---	
Sem. Mk.	457	305	806	905	925	899	---

Teacher 9 N=16 World History Mean IQ=106

Otis	---					
First 6 Wks.	863	---				
Second 6 Wks.	870	893	---			
Third 6 Wks.	848	925	947	---		
Sem. Ex.	764	797	887	817	---	
Sem. Mk.	875	929	980	966	920	---

Teacher 10 Less than 15 students with complete data.

Teacher 11 N=20 Biology Mean IQ=116

Otis	---					
First 6 Wks.	319	---				
Second 6 Wks.	378	700	---			
Third 6 Wks.	224	686	768	---		
Sem. Ex.	539	654	802	822	---	
Sem. Mk.	397	711	793	926	938	---

Teacher 12 N=26 Latin II Mean IQ=121

Otis	---					
First 6 Wks.	522	---				
Second 6 Wks.	524	818	---			
Third 6 Wks.	463	810	753	---		
Sem. Ex.	533	767	809	826	---	
Sem. Mk.	569	879	868	894	944	---

Teacher 13 Data not available.

Teacher 14 N=23 English II Mean IQ=112

Otis	---					
First 6 Wks.	597	---				
Second 6 Wks.	747	690	---			
Third 6 Wks.	678	585	820	---		
Sem. Ex.	705	682	848	709	---	
Sem. Mk.	758	733	869	907	891	---

Teacher 15 N=24 Algebra I Mean IQ=123

Otis	---					
First 6 Wks.	050	---				
Second 6 Wks.	050	928	---			
Third 6 Wks.	-124	768	841	---		
Sem. Ex.	-086	832	841	807	---	
Sem. Mk.	-029	921	972	911	899	---

Teacher 16 Data not available.

Teacher 17 N=24 American History Mean IQ=121

Otis	---					
First 6 Wks.	253	---				
Second 6 Wks.	154	471	---			
Third 6 Wks.	027	223	343	---		
Sem. Ex.	-099	566	045	154	---	
Sem. Mk.	-132	634	503	661	620	---

Teacher 18 Less than 15 students with complete data.

Teacher 19 N=28 Algebra II Mean IQ=113

Otis	---					
First 6 Wks.	254	---				
Second 6 Wks.	166	677	---			
Third 6 Wks.	271	700	708	---		
Sem. Ex.	249	648	692	516	---	
Sem. Mk.	239	878	892	805	839	---

Teacher 20 N=26 Biology Mean IQ=107

Otis	---					
First 6 Wks.	595	---				
Second 6 Wks.	510	750	---			
Third 6 Wks.	546	676	600	---		
Sem. Ex.	573	807	907	714	---	
Sem. Mk.	645	890	879	820	940	---

Teacher 21 Data not available.

Teacher 30 N=20 Advanced Math Mean IQ=119

Otis	---					
First 6 Wks.	107	---				
Second 6 Wks.	421	754	---			
Third 6 Wks.	359	664	759	---		
Sem. Ex.	253	595	597	703	---	
Sem. Mk.	306	863	891	890	827	---

Teacher 31 Less than 15 students with complete data.

Teacher 32 N=25 Biology Mean IQ=109

Otis	---					
First 6 Wks.	427	---				
Second 6 Wks.	319	795	---			
Third 6 Wks.	326	780	771	---		
Sem. Ex.	385	679	746	741	---	
Sem. Mk.	439	864	913	890	886	---

Teacher 33 Less than 15 students with complete data.
Teacher 34 Less than 15 students with complete data.

Teacher 35 N=23 Biology Mean IQ=104

Otis	---					
First 6 Wks.	509	---				
Second 6 Wks.	642	656	---			
Third 6 Wks.	553	616	953	---		
Sem. Ex.	627	726	749	734	---	
Sem. Mk.	615	838	808	807	845	---

Teacher 36 Data not available.

Teacher 37 N=20 Sociology Mean IQ=105

Otis	---					
First 6 Wks.	316	---				
Second 6 Wks.	292	799	---			
Third 6 Wks.	299	877	650	---		
Sem. Ex.	405	741	713	603	---	
Sem. Mk.	406	919	859	845	874	---

Teacher 38 N=18 World History Mean IQ=106

Otis	---					
First 6 Wks.	554	---				
Second 6 Wks.	495	769	---			
Third 6 Wks.	577	721	676	---		
Sem. Ex.	573	687	710	546	---	
Sem. Mk.	647	908	857	769	880	---

Teacher 39 N=20 American History Mean IQ=112

Otis	---					
First 6 Wks.	165	---				
Second 6 Wks.	276	772	---			
Third 6 Wks.	239	882	835	---		
Sem. Ex.	145	860	594	679	---	
Sem. Mk.	158	930	774	858	915	---

Teacher 40 N=19 Algebra I Mean IQ=101

Otis	---					
First 6 Wks.	443	---				
Second 6 Wks.	495	818	---			
Third 6 Wks.	528	864	938	---		
Sem. Ex.	388	890	830	923	---	
Sem. Mk.	502	917	957	979	944	---

Teacher 41 Less than 15 students with complete data.

Teacher 42 N=18 American History Mean IQ=108

Otis	---					
First 6 Wks.	510	---				
Second 6 Wks.	375	171	---			
Third 6 Wks.	369	457	-013	---		
Sem. Ex.	241	495	-183	796	---	
Sem. Mk.	507	761	080	820	875	---

Teacher 43 Less than 15 students with complete data.

Teacher 45 N=23 Spanish I Mean IQ=117

Otis	---					
First 6 Wks.	342	---				
Second 6 Wks.	267	904	---			
Third 6 Wks.	350	850	877	---		
Sem. Ex.	368	811	831	936	---	
Sem. Mk.	351	870	899	985	954	---

Teacher 46 N=22 English II Mean IQ=108

Otis	---					
First 6 Wks.	220	---				
Second 6 Wks.	221	344	---			
Third 6 Wks.	230	750	547	---		
Sem. Ex.	170	678	335	554	---	
Sem. Mk.	287	787	586	780	847	---

Teacher 47 Less than 15 students with complete data.

Teacher 48 N=23 Latin I Mean IQ=112

Otis	---					
First 6 Wks.	470	---				
Second 6 Wks.	576	908	---			
Third 6 Wks.	515	856	820	---		
Sem. Ex.	473	865	824	870	---	
Sem. Mk.	559	956	956	899	868	---